



(12) **United States Patent**
Pugachev et al.

(10) **Patent No.:** **US 9,217,158 B2**
(45) **Date of Patent:** **Dec. 22, 2015**

(54) **REPLICATION-DEFECTIVE FLAVIVIRUS VACCINES AND VACCINE VECTORS**
(71) Applicants: **Konstantin V. Pugachev**, Natick, MA (US); **Maryann Giel-Moloney**, Brighton, MA (US); **Harold Kleanthous**, Westford, MA (US); **Mark Parrington**, Bradford (CA); **Linong Zhang**, Maple (CA)

2006/0159704 A1 7/2006 Bonaldo et al.
2006/0204523 A1 9/2006 Khromykh et al.
2007/0207166 A1* 9/2007 Nabel et al. 424/199.1
2007/0218078 A1 9/2007 Clarke et al.
2007/0249032 A1 10/2007 Pang et al.
2008/0175862 A1 7/2008 Pugachev et al.
2011/0135686 A1 6/2011 Pugachev et al.

(72) Inventors: **Konstantin V. Pugachev**, Natick, MA (US); **Maryann Giel-Moloney**, Brighton, MA (US); **Harold Kleanthous**, Westford, MA (US); **Mark Parrington**, Bradford (CA); **Linong Zhang**, Maple (CA)

FOREIGN PATENT DOCUMENTS

EP 0117657 A1 9/1984
WO WO-98/37911 A1 9/1998
WO WO-99/28487 A1 6/1999
WO WO-02/102828 A2 12/2002
WO WO-03/046189 A1 6/2003
WO WO-2006/086838 A1 8/2006
WO WO-2007/098267 A2 8/2007
WO WO-2008/137163 A1 11/2008
WO WO 2009/114207 A2* 9/2009
WO WO-2010/107847 A1 9/2010

(73) Assignee: **Sanofi Pasteur Biologics, LLC**, Cambridge, MA (US)

OTHER PUBLICATIONS

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 139 days.

Jamieson et al., *Lancet*, 2009, 374:451-458.*
Shustov et al., *Journal of Virology*, available online Aug. 22, 2007, 81(21):11737-11748.*
International Search Report and Written Opinion from PCT/US2013/024495, mailed Apr. 11, 2013 (11 pages).
International Preliminary Report on Patentability for PCT/US2013/024495, dated Aug. 5, 2014 (8 pages).
Bonaldo et al., "The yellow fever 17D vaccine virus as a vector for the expression of foreign proteins: development of new live flavivirus vaccines," *Mem Inst Oswaldo Cruz. Suppl* 1:215-23 (2000).
Bredendiek et al., "A recombinant Yellow Fever 17D vaccine expressing Lassa virus glycoproteins," *Virology*. 345(2):299-304 (2006).
Hsu et al., "Efficacy of adenovirus-vectored respiratory syncytial virus vaccines in a new ferret model," *Vaccine*. 12(7):607-12 (1994).
Li et al., "Chimeric influenza virus induces neutralizing antibodies and cytotoxic T cells against human immunodeficiency virus type 1," *J Virol*. 67(11):6659-66 (1993).
Martinez-Sobrido et al., "Protection against respiratory syncytial virus by a recombinant Newcastle disease virus vector," *J Virol*. 80(3):1130-9 (2006).
McAllister et al., "Recombinant yellow fever viruses are effective therapeutic vaccines for treatment of murine experimental solid tumors and pulmonary metastases," *J Virol*. 74(19):9197-205 (2000).
Pang et al., "Development of dengue virus replicons expressing HIV-1 gp120 and other heterologous genes: a potential future tool for dual vaccination against dengue virus and HIV," *BMC Microbiol*. 1(28):1-9 (2001).
Shustov et al., "Production of pseudoinfectious yellow fever virus with a two-component genome," *J Virol*. 81(21):11737-48 (2007).
Extended European Search Report for European Patent Application No. 097020269.1, dated Mar. 26, 2012 (9 pages).

(21) Appl. No.: **13/633,436**

(22) Filed: **Oct. 2, 2012**

(65) **Prior Publication Data**
US 2013/0243812 A1 Sep. 19, 2013

Related U.S. Application Data

(63) Continuation of application No. 13/364,187, filed on Feb. 1, 2012, now abandoned, which is a continuation-in-part of application No. 12/922,513, filed as application No. PCT/US2009/001666 on Mar. 16, 2009, now Pat. No. 8,815,564.

(60) Provisional application No. 61/092,814, filed on Aug. 29, 2008, provisional application No. 61/069,451, filed on Mar. 14, 2008.

(51) **Int. Cl.**
A61K 39/12 (2006.01)
C12N 15/86 (2006.01)
A61K 39/295 (2006.01)

(52) **U.S. Cl.**
CPC **C12N 15/86** (2013.01); **A61K 39/295** (2013.01); **C12N 2770/24121** (2013.01); **C12N 2770/24134** (2013.01)

(58) **Field of Classification Search**
None
See application file for complete search history.

(56) **References Cited**
U.S. PATENT DOCUMENTS

4,393,201 A 7/1983 Curtis et al.
5,830,477 A 11/1998 Lathe et al.
6,486,135 B1 11/2002 Li et al.
6,696,281 B1 2/2004 Monath et al.
2001/0046499 A1 11/2001 Kantor et al.
2004/0005542 A1 1/2004 Krempl et al.
2004/0223979 A1 11/2004 Chambers et al.

(Continued)

Primary Examiner — Stacy B Chen
(74) *Attorney, Agent, or Firm* — Clark & Elbing LLP; Susan M. Michaud

(57) **ABSTRACT**

This invention provides replication-defective flavivirus vaccines and vaccine vectors, and corresponding compositions and methods.

8 Claims, 37 Drawing Sheets

(56)

References Cited

OTHER PUBLICATIONS

International Preliminary Report on Patentability for International Application No. PCT/US09/01666, dated Apr. 25, 2011 (date of completion of report) and Jun. 20, 2011 (date of mailing of report). 10 pages.

International Search Report for International Patent Application No. PCT/US2009/001666, dated Aug. 12, 2009 (date of completion of search) and Sep. 1, 2009 (date of mailing of report). 3 pages.

International Search Report for International Patent Application No. PCT/US2010/027552, dated May 20, 2010 (date of completion of search) and Jun. 15, 2010 (date of mailing of report). 4 pages.

Search Report and Written Opinion issued by the Hungarian Intellectual Property Office for Singapore Patent Application No. 201006578-7, dated Jul. 19, 2012 (19 pages).

Written Opinion for International Patent Application No. PCT/US2009/001666, dated Aug. 12, 2009 (date of completion of opinion) and Sep. 1, 2009 (date of mailing of opinion). 9 pages.

Written Opinion for International Patent Application No. PCT/US2010/027552, dated May 20, 2010 (date of completion of opinion) and Jun. 15, 2010 (date of mailing of opinion). 5 pages.

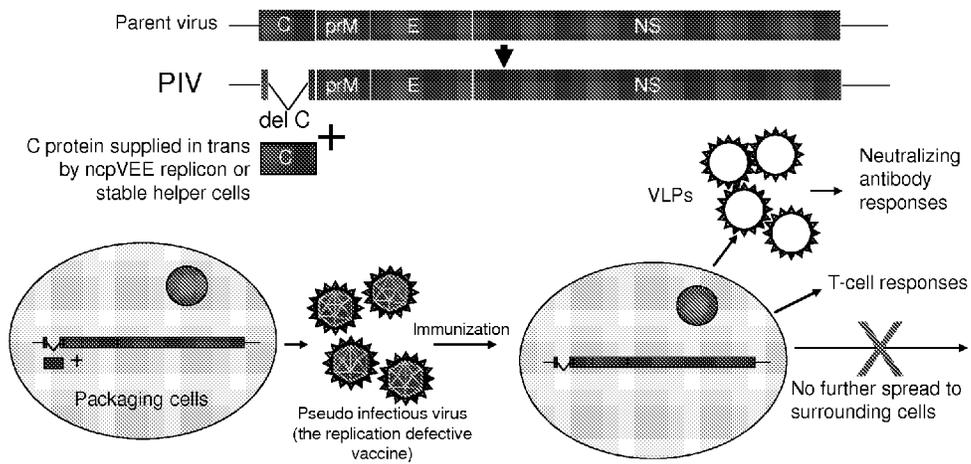
Scherret et al., "The relationships between West Nile and Kunjin viruses," *Emerg Infect Dis.* 7(4):697-705 (2001).

Weyer et al., "Generation and evaluation of a recombinant modified vaccinia virus Ankara vaccine for rabies," *Vaccine.* 25(21):4213-22 (2007).

Search Report and Written Opinion for Singapore Patent Application No. 2013018320, dated May 6, 2015 (22 pages).

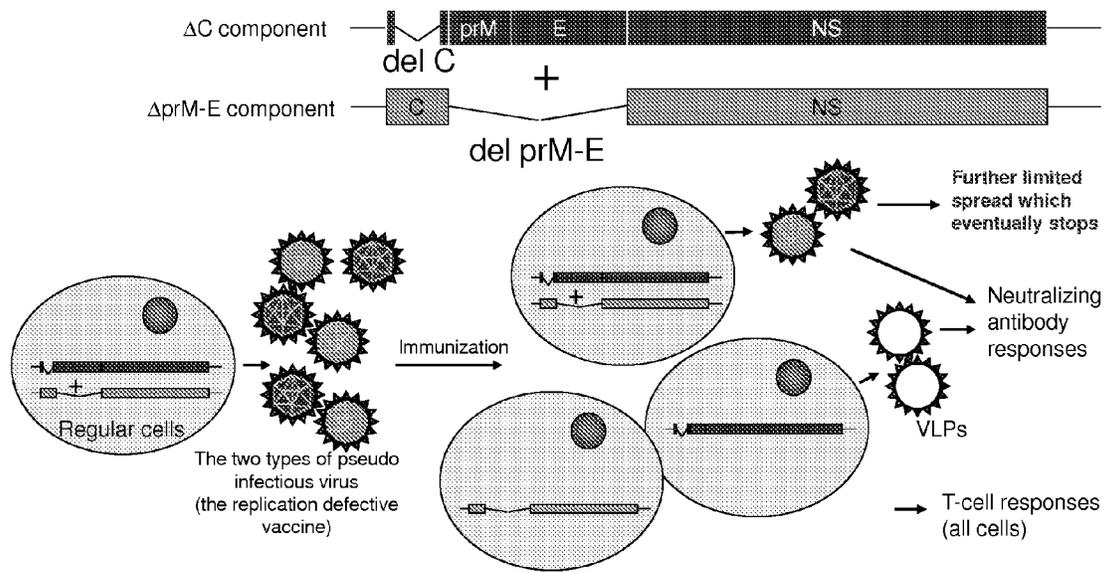
* cited by examiner

Fig. 1. Principle of single-component PIV (s-PIV; single-round replication in vivo)



For recombinant vaccines: foreign immunogen inserted in place of the ΔC deletion, or elsewhere, e.g. intergenically etc.

Fig. 2. Principle of two-component PIV (d-PIV; limited spread in vivo)



For recombinant vaccines: foreign immunogen inserted in place of ΔC and/or ΔprM-E, or elsewhere.

Fig. 3. Immunogenicity/efficacy: general experiment design (mice)

Experiment design in mice

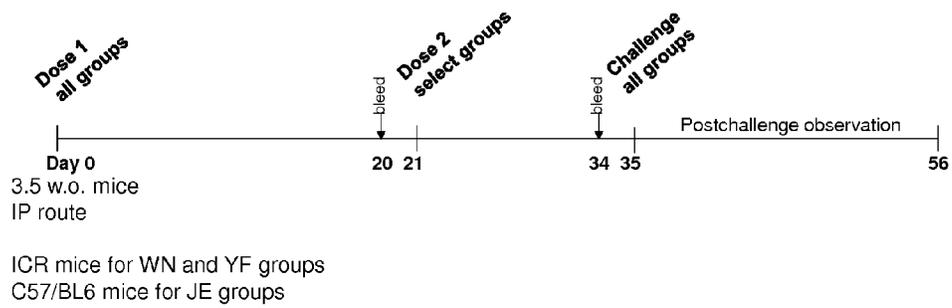
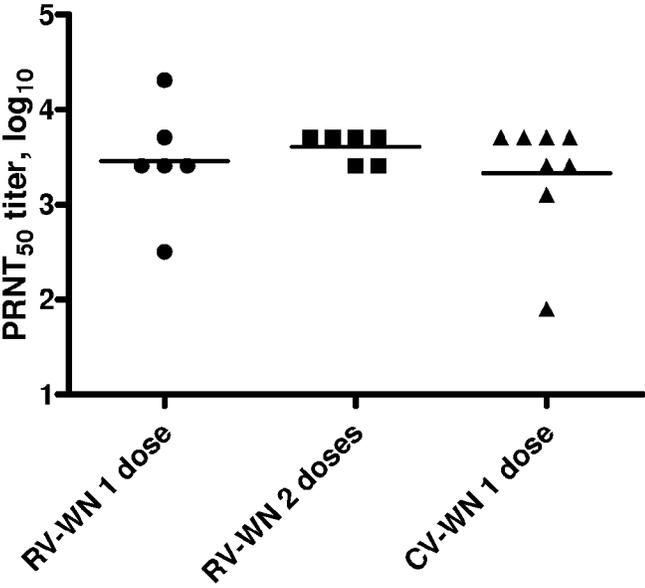


Fig. 4. PIV-WN induces uniform humoral immunity in mice



ICR mice, IP inoculation

Fig. 5. PIV-YF and PIV-WN protect hamsters against post-challenge viremia and morbidity

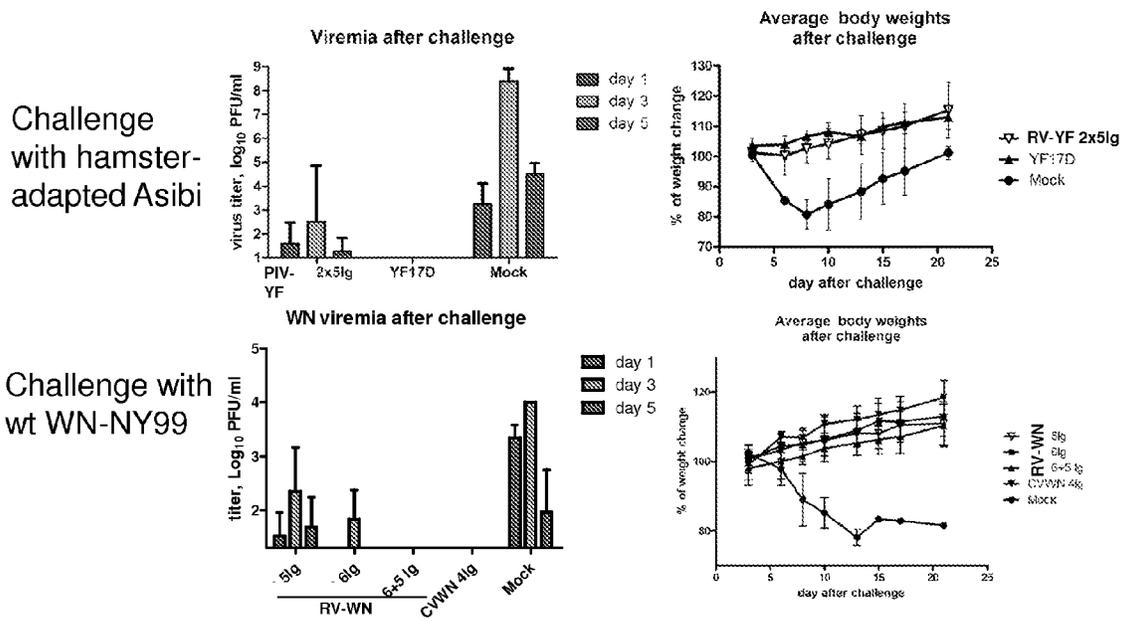
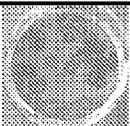
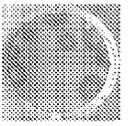


Fig. 6. YF/TBE viruses

| prM-E genes | P0 virus titer log ₁₀ PFU/ml | P1 virus titer log ₁₀ PFU/ml | Immuno-staining αRSSE mHIAF | |
|--------------------------|--|--|--|------|
| TBEV prM-SP (p42) | 7.6 | 7.9 |  | <p42 |
| Hypr WNV prM-SP (p45) | 7.1 | 7.4 | | p45> |
| Hypr + dC2 "CΔ3aa" (p59) | 5.6 | 6.5 |  | p59> |
| LGT/E5 (p43) | 7.8 | 8.1 |  | <p43 |

NOTES:

- p42, p45, p59, and p43 are designations of plasmids
- plaque morphology for p59-derived chimera was determined in a separate titration experiment (not shown; result of immunofluorescence assay shown)

Fig. 7. PIV-WNV/TBE constructs

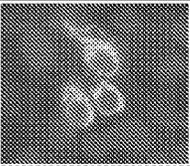
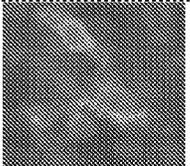
| prM-E genes | P0 titer log ₁₀ FFU/ml | | P1 titer log ₁₀ FFU/ml | | Immuno-staining αRSSE mHIAF |
|-----------------------------|--------------------------------------|--------------------|--------------------------------------|--------------------|---|
| | C helper cells | CprME helper cells | C helper cells | CprME helper cells | |
| Hypr (TBEV signal) (p39) | 7.2 | 6.7 | 6.9 | 77.1 |  |
| Hypr (WNV signal) (p40) | 6.7 | 6.0 | 5.9 | 66.9 |  |

Fig. 8. Replication kinetics of live YF/TBE and replication defective PIV-WNV/TBE variants in cell substrates

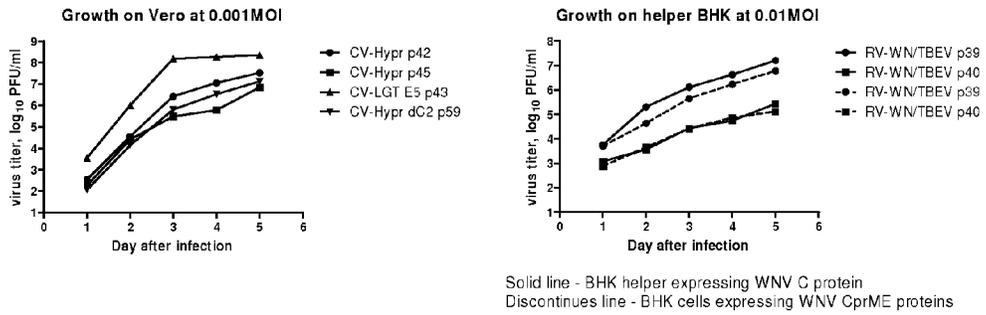


Fig. 9. Survival of mice inoculated IC with PIV-TBE and YF/TBE constructs in the neurovirulence test (3.5 week-old ICR mice)

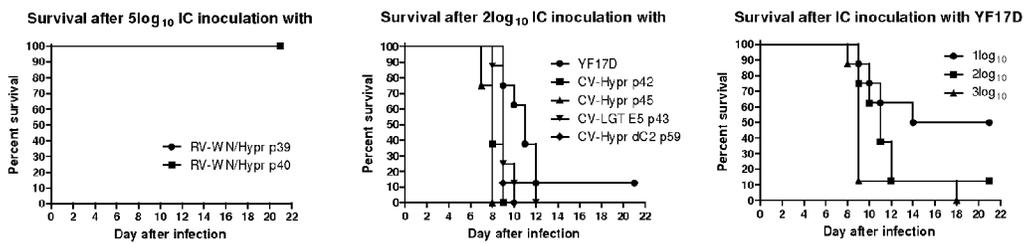


Fig. 10. Survival of mice inoculated IP with PIV-TBE and YF/TBE constructs in a neuroinvasiveness test (3.5 week-old ICR mice)

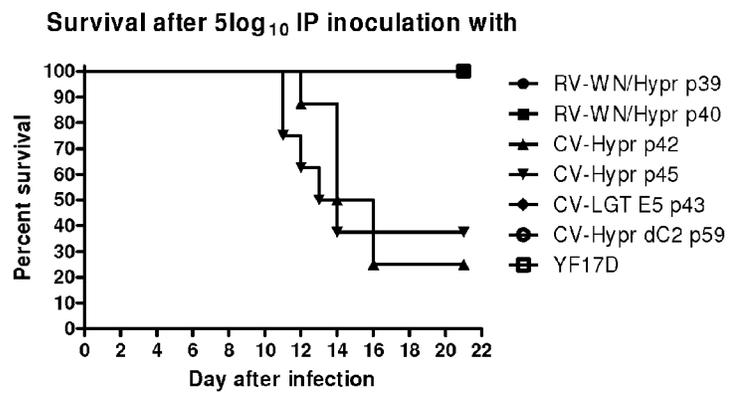


Fig. 11. Post-TBE-challenge morbidity (weight loss); day 9 post-challenge

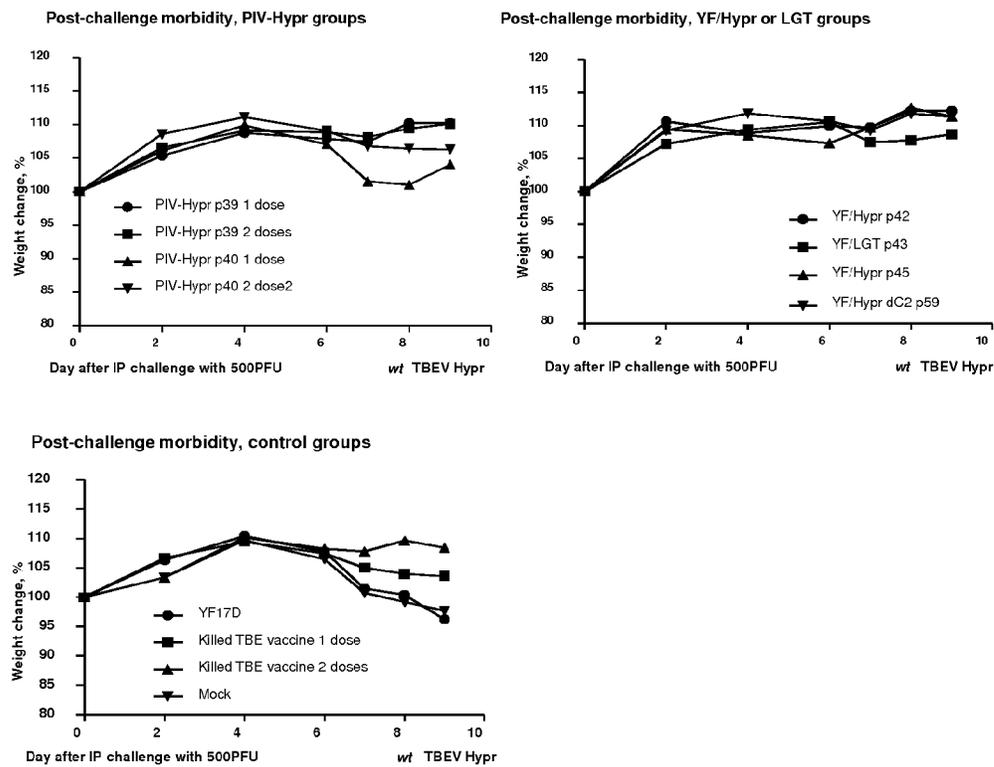


Fig. 12. Examples of PIV constructs expressing foreign antigens (rabies G)

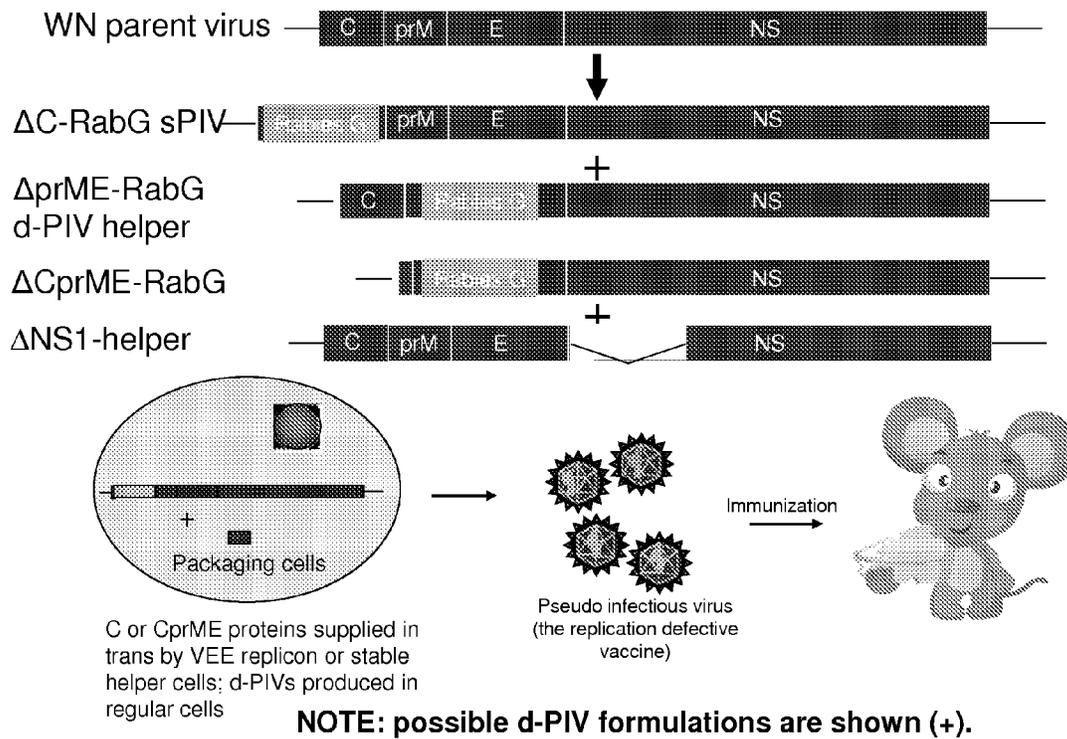


Fig. 13. Schematic representation of insertion designs resulting in viable/expressing constructs (exemplified by rabies G)

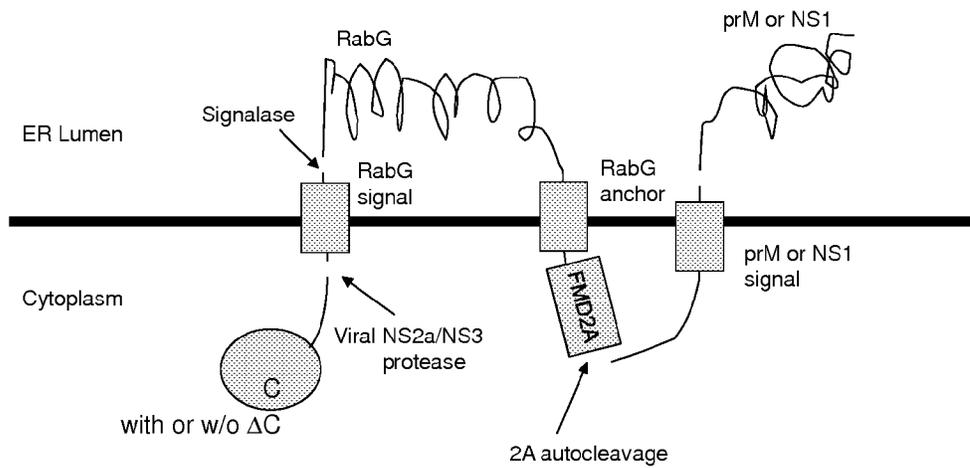


Fig. 14. PIV-WN/Rabies G PIVs: immunofluorescence of transfected cells and growth curves after transfection

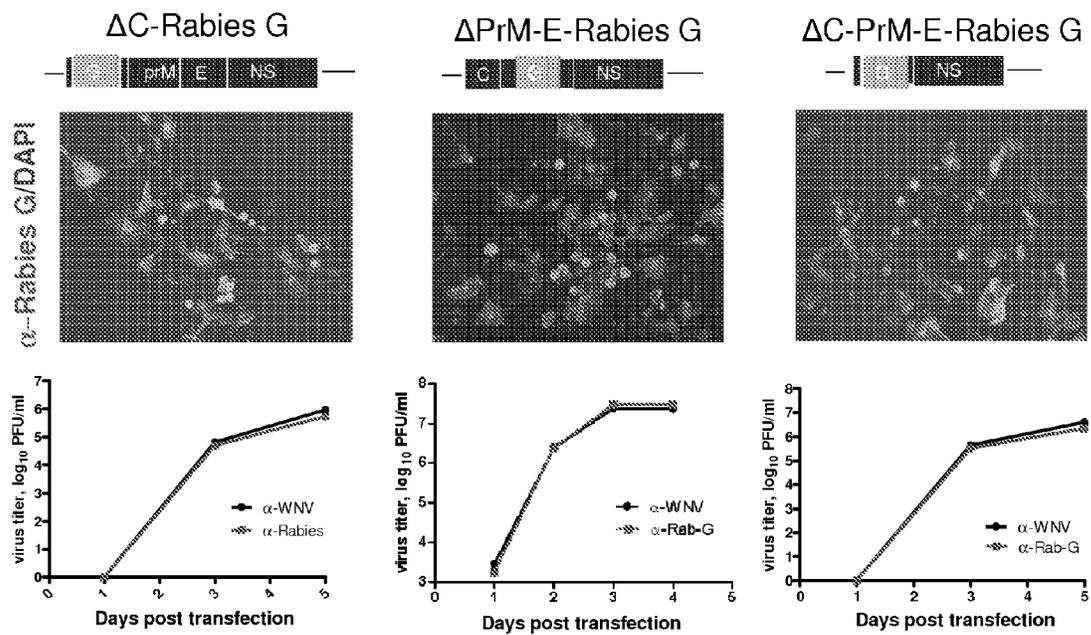
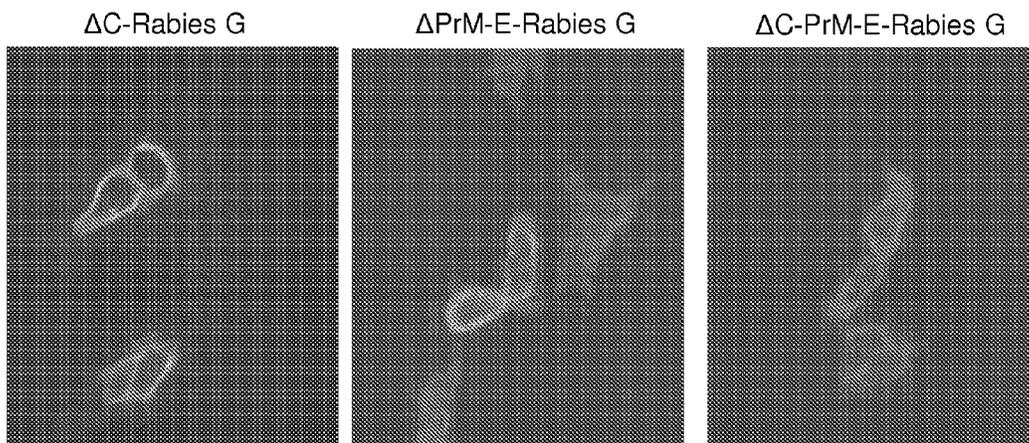


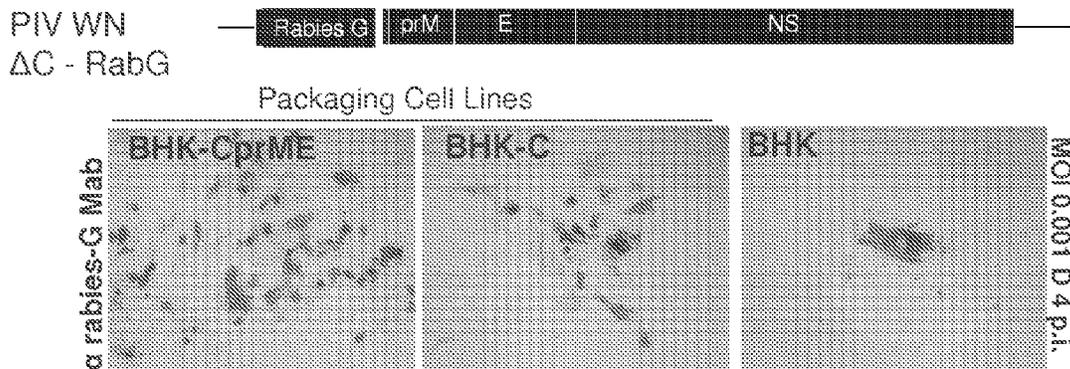
Fig. 15. Efficient expression of RabG on the plasma membrane of Vero cells

MOI 0.1, day 2 post infection, 4% PFA Fixed non-permeabilized



1Ab (Abcam) Anti Rabies 1:500 2-3hrs RT, 2Ab Anti-Mouse IgG 1:1000 1hr RT
Images are 40X and exposure time is 40ms

Fig. 16. PIV - Rabies G spread in helper cells, no spread in naïve cells



PIVs spread in helper cells and not naïve cells.
No infectious material in supernatants from BHK cells.

Similarly, VSV G was expressed, and no spread in regular cells (Vero, BHK) was observed and no infectious material detected in the supernatants

Fig. 17. Stability of rabies G protein gene in PIV-WN vectors
 Passages in BHK-C-prM-E helper cells, MOI 0.1; titration in Vero by immunostaining

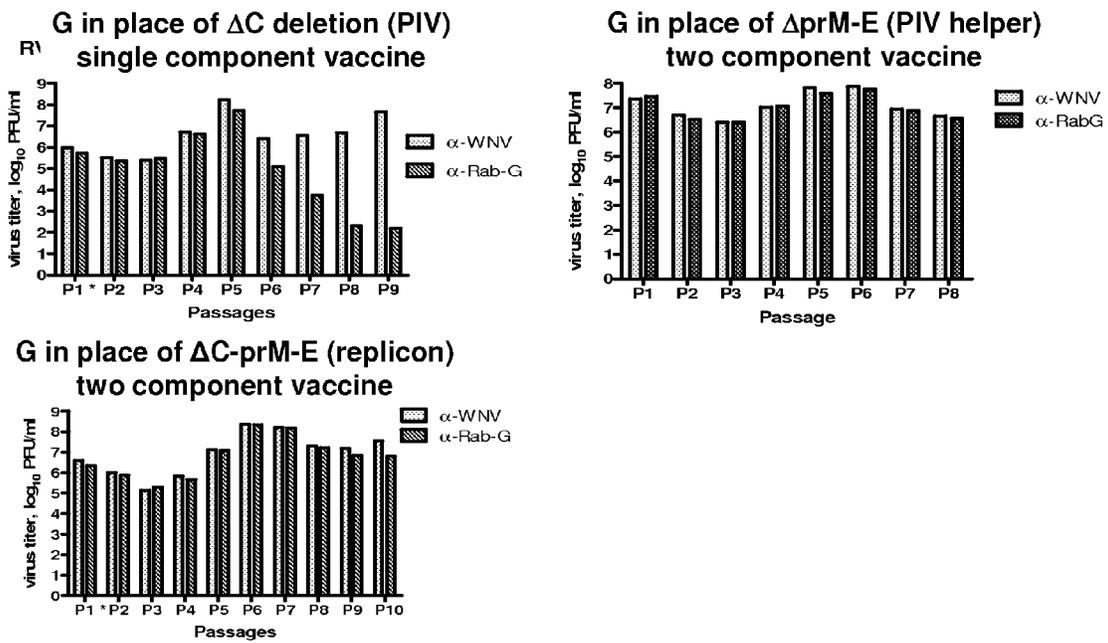


Fig. 18. Comparison of spread in Vero cells of single-component vs. two-component PIV-Rabies variants

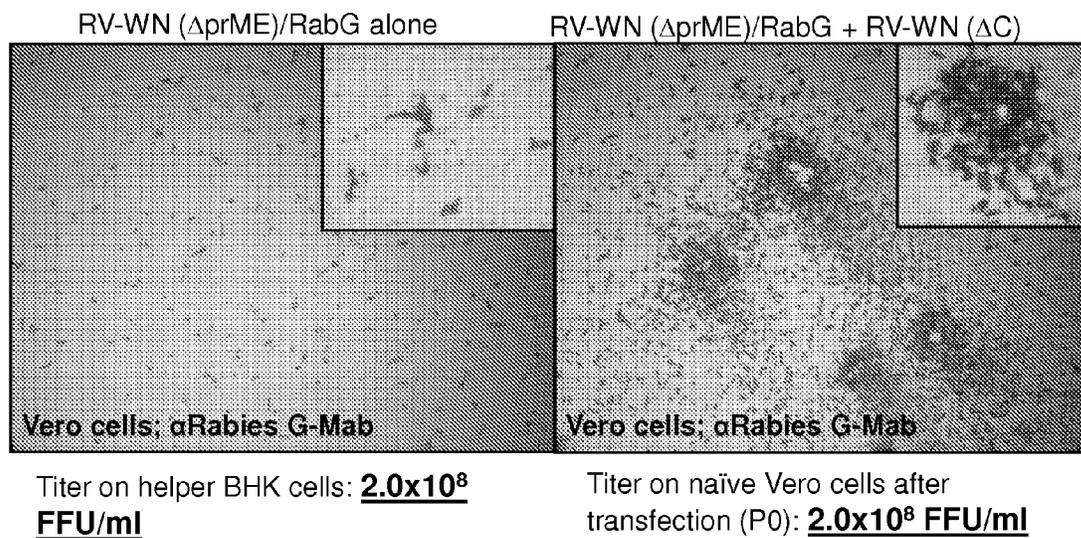
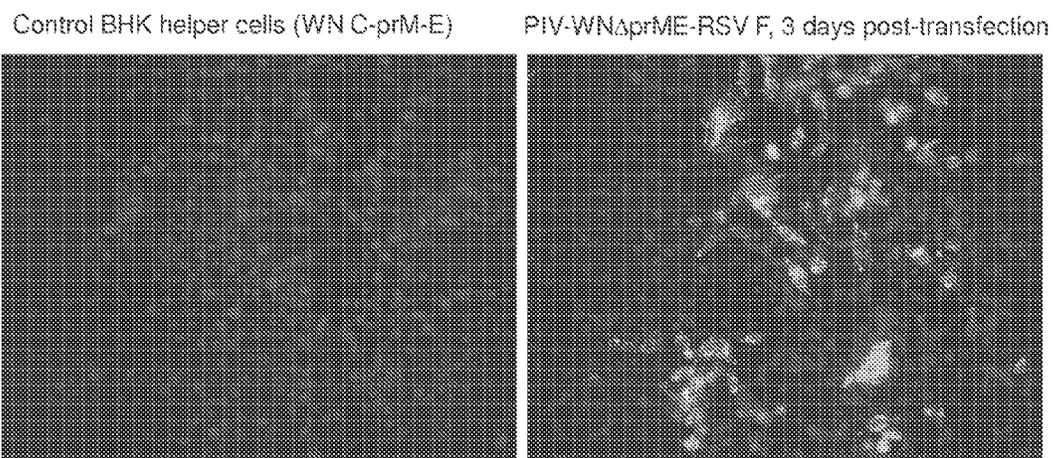
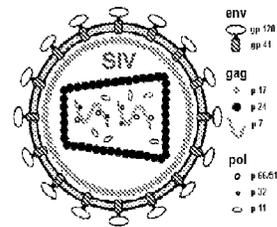


Fig. 19. Example of expression of full-length RSV F protein (strain A2): immunostaining of helper cells after transfection



Cells stained with anti-RSVF Mab, DAPI

Fig. 20. PIV (WN) SIV Constructs



Synthesized genes: sequences optimized for repeats

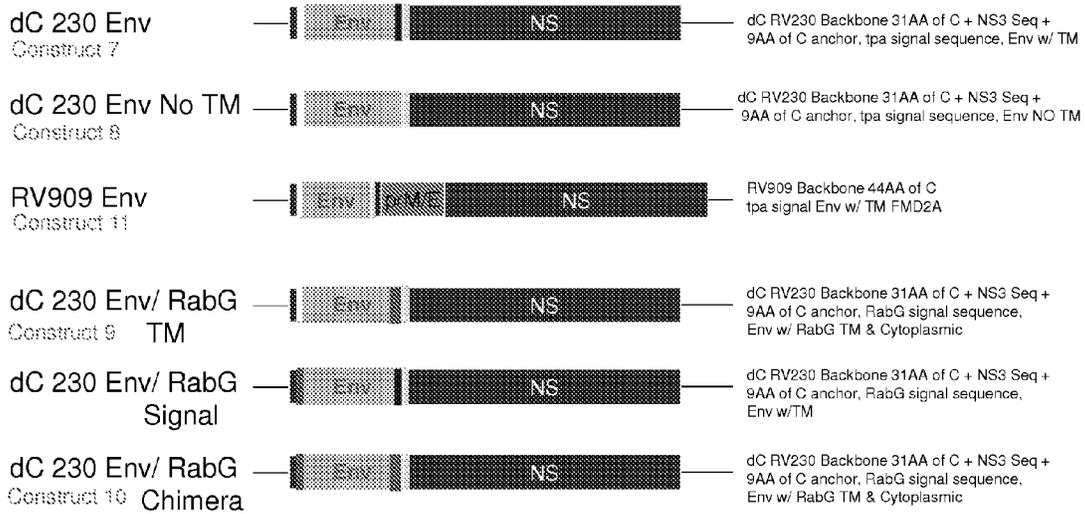


RV WN Vectors: Δ C-prM-E, Δ prM-E, and Δ C

Construct Designs:



Fig. 21. Env Construct Designs



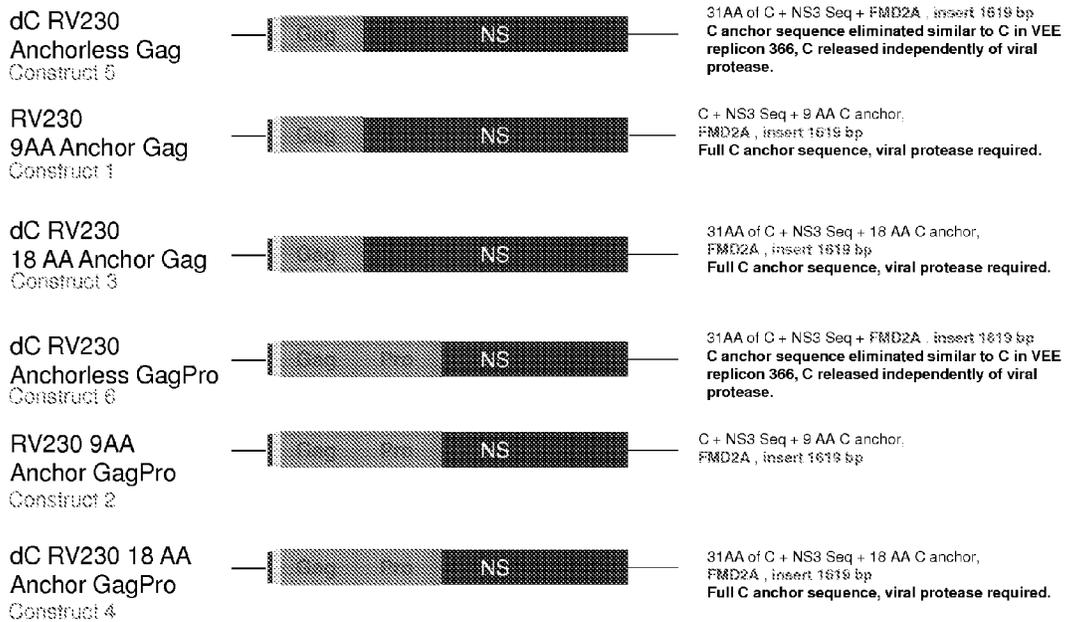
Construct #s indicate design sequences in Appendix 1

Env RabG Signal= Env with RabG Signal and Env TM

Env RabG TM= Env with tpa signal and RabG TM

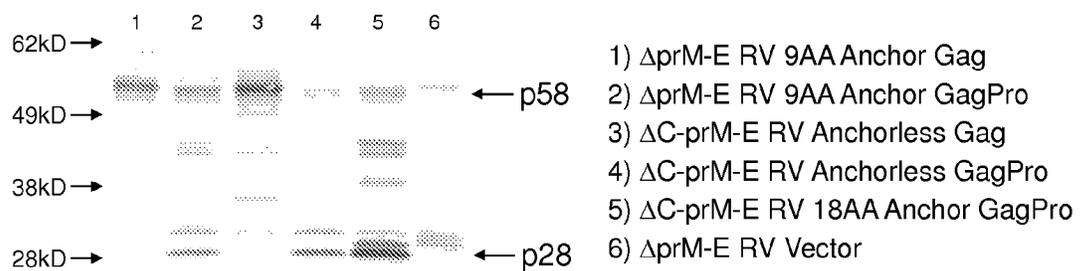
Env RabG Chimera= Env with rabG signal and TM

Fig. 22. Gag Construct Designs



Construct #s indicate design sequences in Appendix 1

Fig. 23. Gag Western demonstrating correct polyprotein processing



BHK C-prM-E packaging cell lysate 3 days post infection. 1Ab Anti-SIVp28, 2Ab Anti Rabbit IgG

Fig. 24. Immunostaining of RV-SIV Gag infected naïve Vero cells.

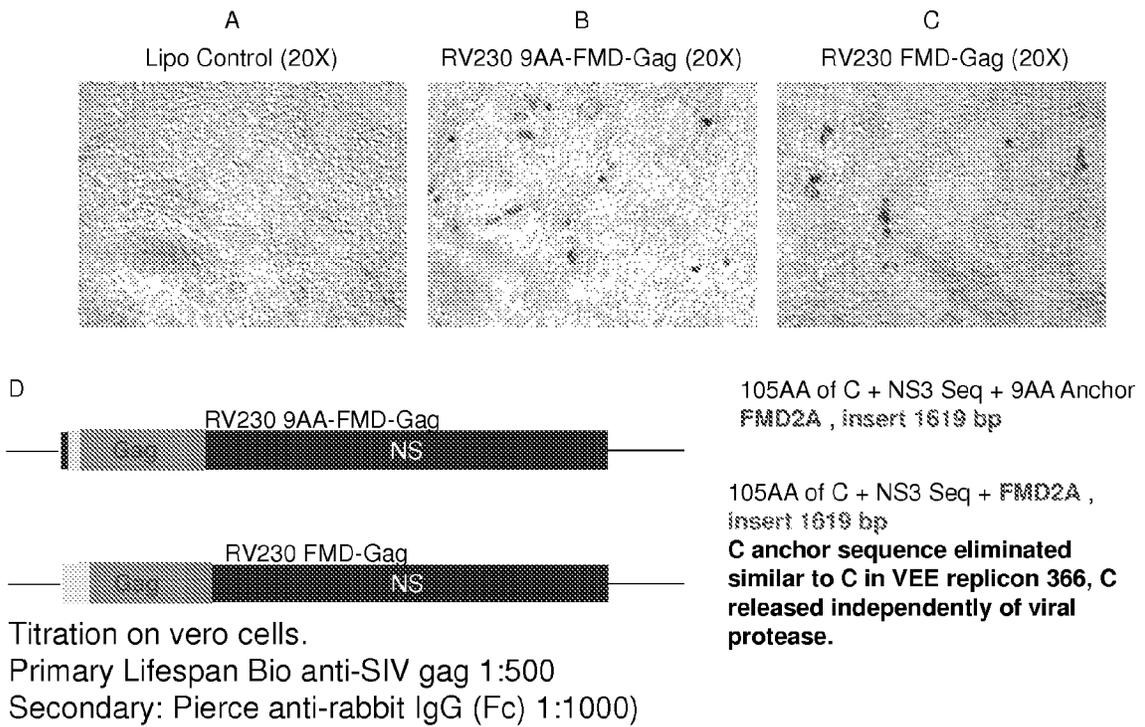
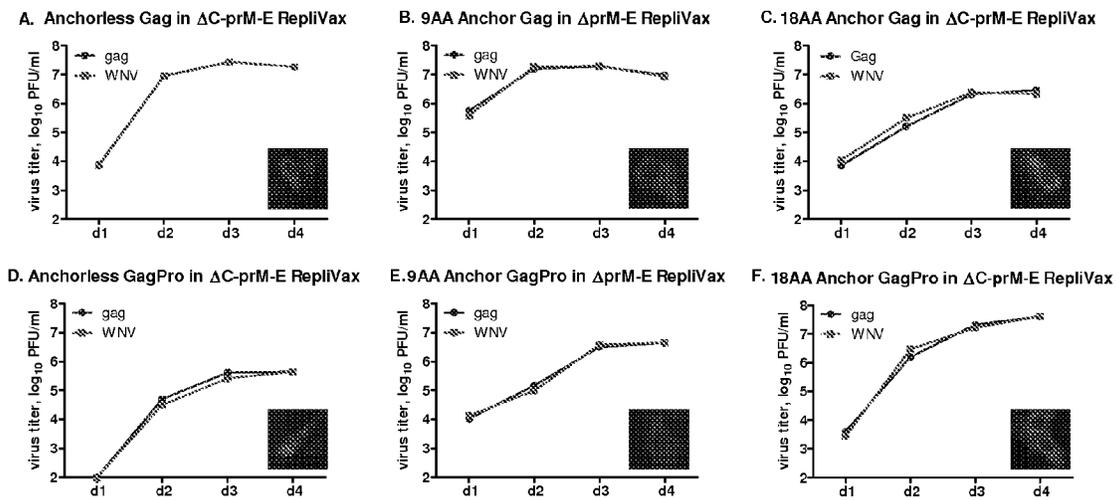
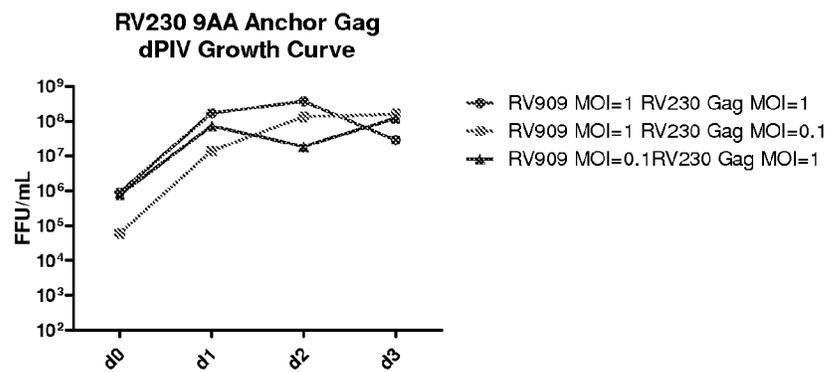


Fig. 25. RV-SIV Gag & GagPro Growth Curves. PIVs recovered in BHK-C-prM-E helper cells; titration in Vero by immunostaining.



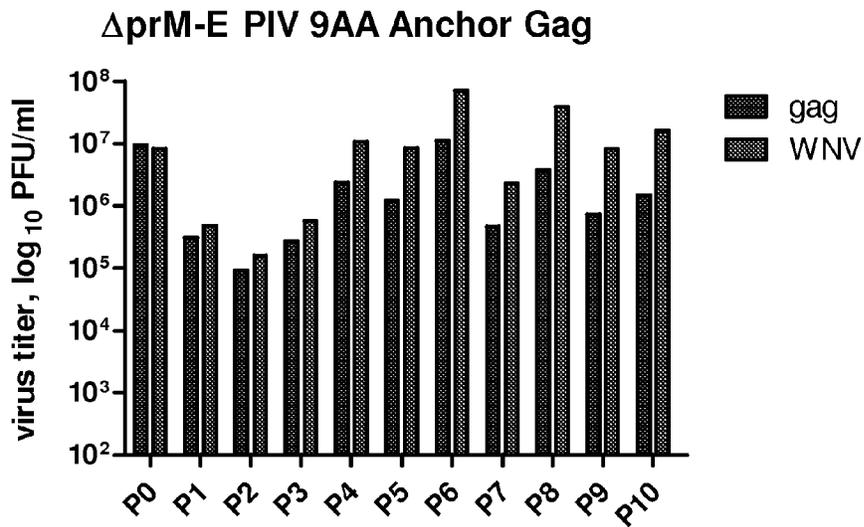
- Similar high titers achieved for WN and SIV antigens post transfection

Fig. 26. d-PIV Gag Growth Curves in naïve Vero.



Vero cells coinfecting with RV230 9AA Anchor Gag and RV230 at various MOIs. Gag specific titers are shown.

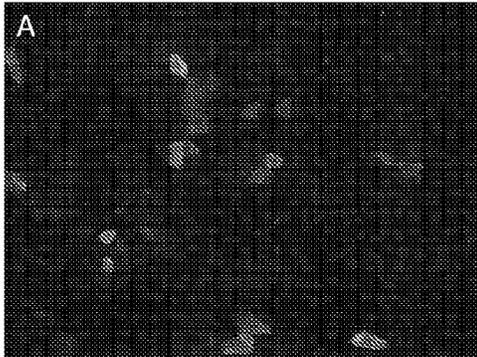
Fig. 27. Stability of the large Gag insert of RV-SIV Gag after 10 serial passages in helper cells, as demonstrated by titration using anti-WN and anti-Gag antibodies.



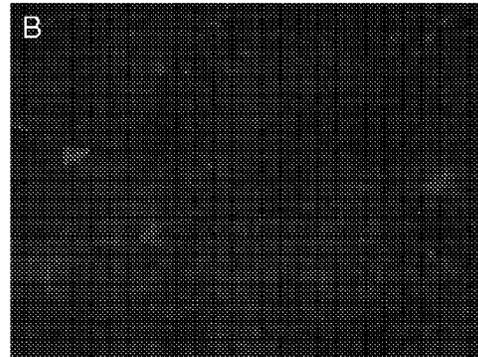
Serial passages of gag constructs on BHK363 cells (P10-P20) @ MOI=0.1 relative to gag titers

Fig. 28. Expression of SIV Env, and better surface presentation using RabG TM.

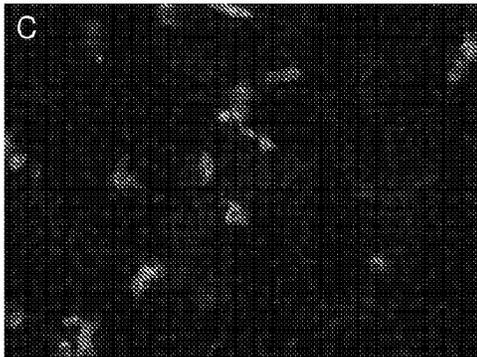
dC230 Env RabG Chimera #7 4%PFA Fix (20X)



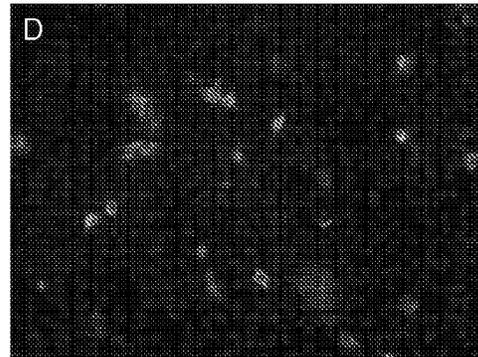
dC230 Env #12 4% PFA Fix (20X)



dC230 Env RabG Chimera #7 Methanol Fix (20X)



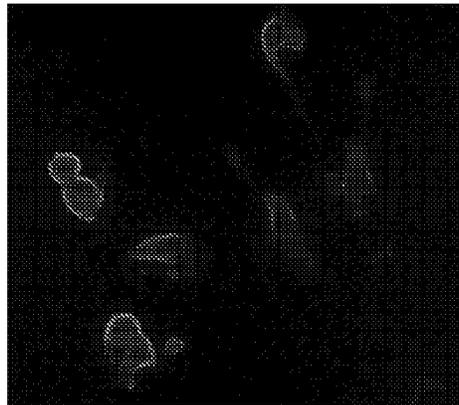
dC230 Env #12 Methanol Fix (20X)



Stained with rabbit polyclonal specific for Env 1:500. Anti rabbit alexa 488 secondary, followed by DAPI stain

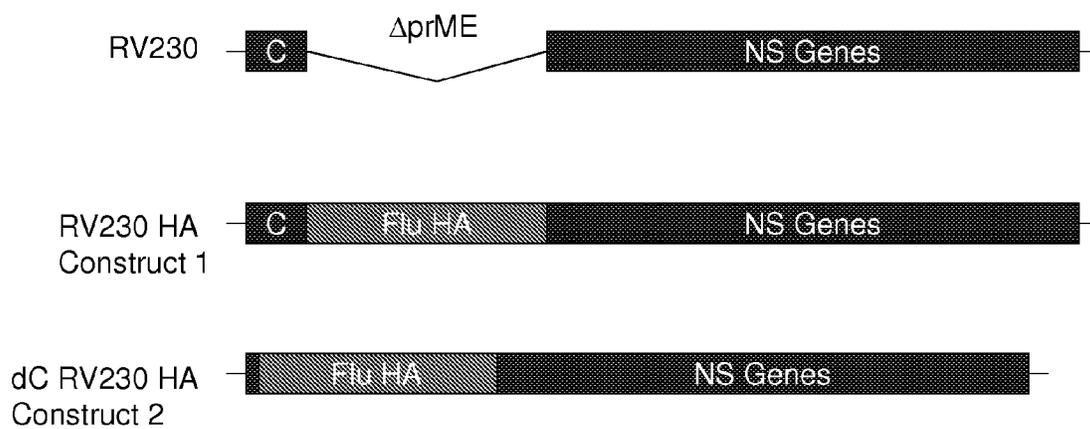
Fig. 29. SIV Env on the surface of PIV-SIV Env/RabG TM infected Vero cells

dC RV230 Env RabG TM



20X Magnification of Vero cells infected at a MOI of ~0.3 and fixed with 4% PFA
Primary: Genway Bio Anti-SIV Env rabbit poly 1:500.
Secondary: Anti rabbit IgG Alexa 488 1:1000

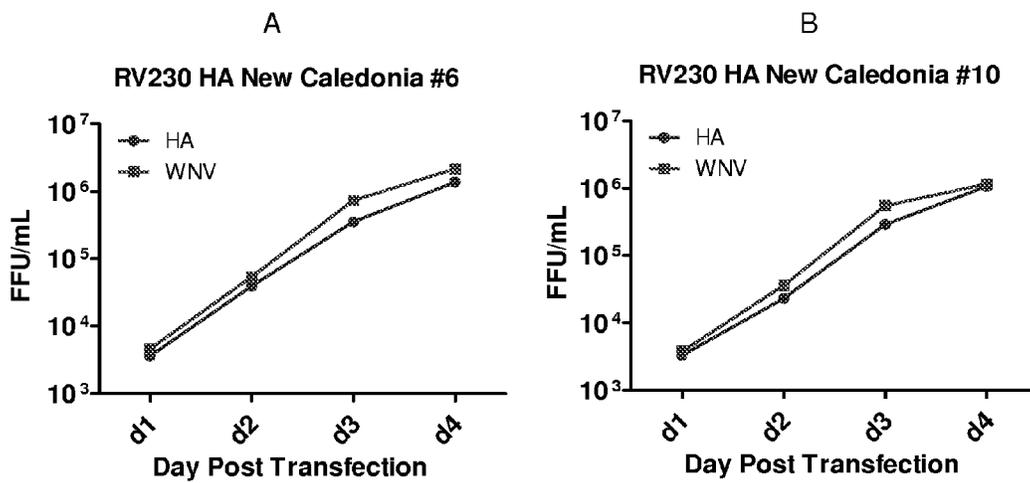
Fig. 30. PIV-flu HA Construct Designs



Blue blocks denote WNV backbone
 Green blocks denote inserted HA gene

Construct #s indicate sequences in Appendix 3

Fig. 31. RV230 HA New Caledonia P0 Growth Curves.



BHK 363 cells transfected at P14 with RNA from RV230 HA New Caledonia clones 6 and 10

Fig. 32. RV230 & Δ C RV230 HA New Caledonia Growth Curves.

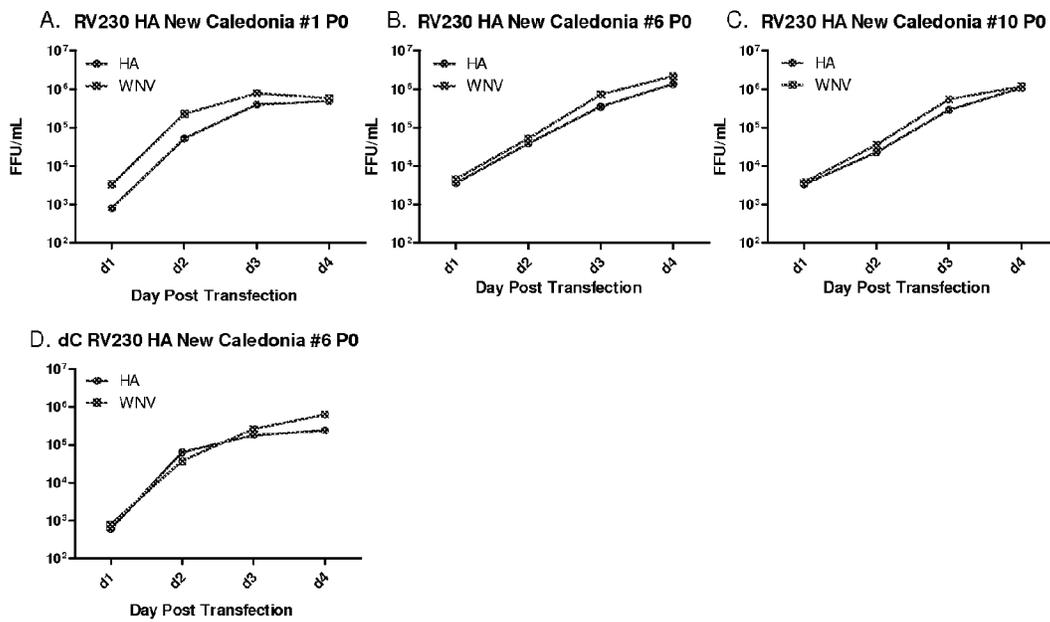
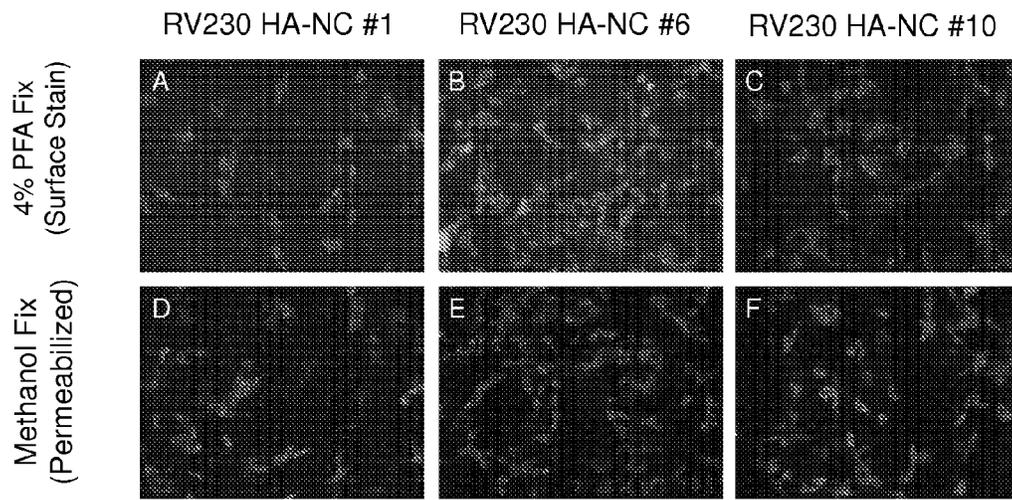


Fig. 33. RV230 HA New Caledonia IF



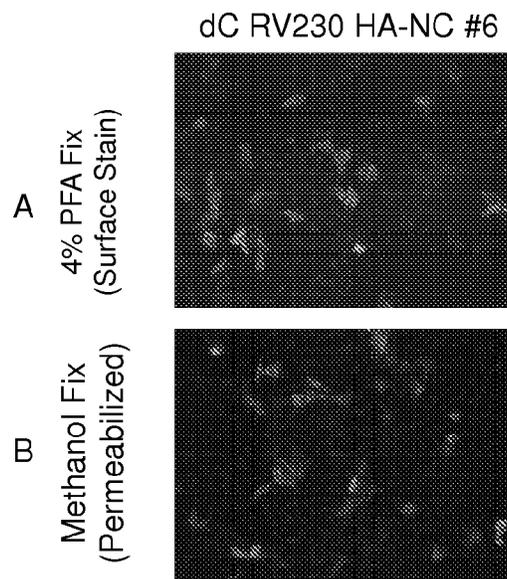
C179 & IT-003-001M2 antibodies pooled and used at final dilution of 1:1000

Pictures taken at 20X Magnification

Exposure Times: PFA Fix- HA 500 ms, DAPI 200 ms

Methanol Fix: HA 200 ms, DAPI 100 ms

Fig. 34. dC RV230 HA New Caledonia IF



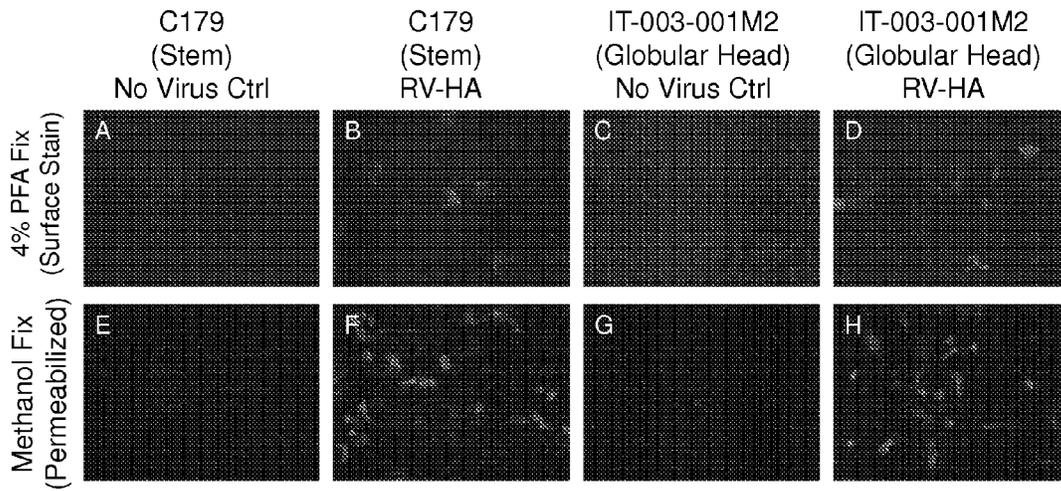
C179 & IT-003-001M2 antibodies pooled and used at final dilution of 1:1000

Pictures taken at 20X Magnification

Exposure Times: PFA Fix- HA 500 ms, DAPI 200 ms

Methanol Fix: HA 200 ms, DAPI 100 ms

Fig. 35. Control IF of uninfected cells (see no virus controls).



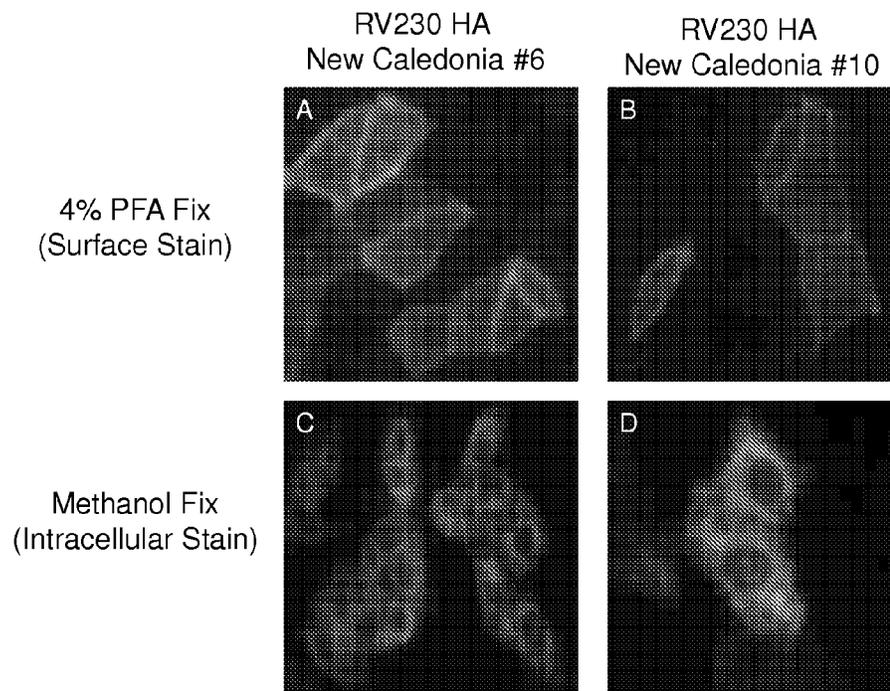
C179 & IT-003-001M2 antibodies used at 1:500

Pictures taken at 20X Magnification

Exposure Times: PFA Fix- HA 500 ms, DAPI 200 ms

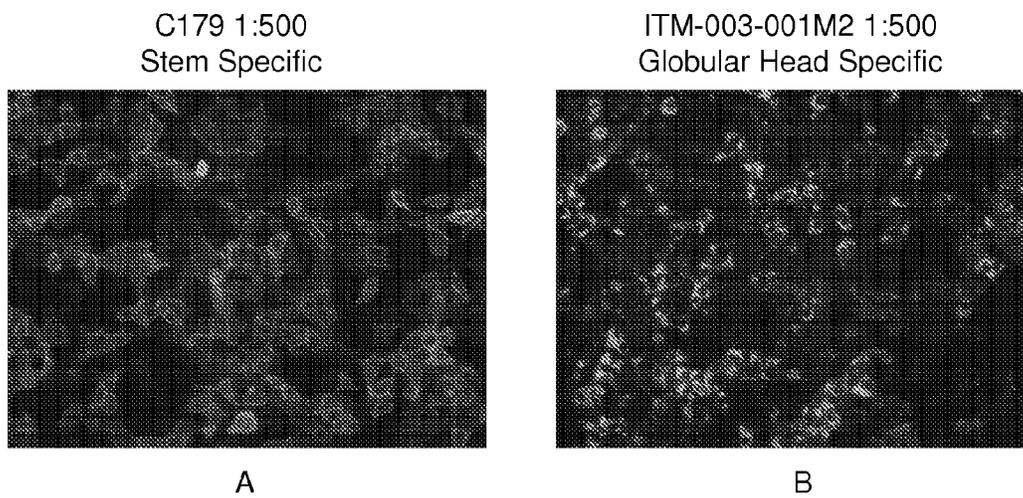
Methanol Fix: HA 200 ms, DAPI 100 ms

Fig. 36. RV-HA New Caledonia IF Staining



Vero cell staining 48 hr post infection, 20X magnification
4% PFA fix 800 ms exposure, methanol fix 300 ms exposure, 20X Magnification
HA stained with 1:500 final pool of C179 (stem specific) and ITM-003-001M2 (globular head specific) mAbs

Fig. 37. Efficient staining of RV230-HA infected Vero cells separately by anti-HA stem and anti-HA globular head antibodies.



Vero cells 48 hr pi with RV230 HA New Caledonia,
4% PFA fixation, 800 ms exposure

REPLICATION-DEFECTIVE FLAVIVIRUS VACCINES AND VACCINE VECTORS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of U.S. Ser. No. 13/364, 187, filed Feb. 1, 2012 (abandoned), which is a continuation in part of U.S. Ser. No. 12/922,513, filed Sep. 14, 2010 (U.S. Pat. No. 8,815,564), which is the U.S. national stage filing under 35 U.S.C. §371 of international application PCT/US2009/001666, filed Mar. 16, 2009, which claims benefit of Provisional Application Nos. 61/069,451, filed Mar. 14, 2008 and 61/092,814, filed Aug. 29, 2008. The prior applications are incorporated herein by reference.

Sequence Listing

The instant application contains a Sequence Listing which has been filed electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Sept. 18, 2015, is named 06132-152003_SL.txt and is 477,191 bytes in size.

FIELD OF THE INVENTION

This invention relates to replication-defective flavivirus vaccines and vaccine vectors, and corresponding compositions and methods.

BACKGROUND OF THE INVENTION

Flaviviruses are distributed worldwide and represent a global public health problem. Flaviviruses also have a significant impact as veterinary pathogens. Flavivirus pathogens include yellow fever (YF), dengue types 1-4 (DEN 1-4), Japanese encephalitis (JE), West Nile (WN), tick-borne encephalitis (TBE), and other viruses from the TBE serocomplex, such as Kyasanur Forest disease (KFD) and Omsk hemorrhagic fever (OHF) viruses. Vaccines against YF [live attenuated vaccine (LAV) strain 17D], JE [inactivated vaccines (INV) and LAV], and TBE (INV) are available. No licensed human vaccines are currently available against DEN and WN. Veterinary vaccines have been in use including, for example, vaccines against WN in horses (INV, recombinant and live chimeric vaccines), JE (INV and LAV) to prevent encephalitis in horses and stillbirth in pigs in Asia, louping ill flavivirus (INV) to prevent neurologic disease in sheep in the UK, and TBE (INV) used in farm animals in Czech Republic (INV) (Monath and Heinz, *Flaviviruses*, in Fields et al. Eds., *Fields Virology*, 3rd Edition, Philadelphia, New York, Lippincott-Raven Publishers, 1996, pp. 961-1034).

Tick-borne encephalitis (TBE) is the most important tick-borne viral disease of humans. It is endemic in parts of Europe and Northern Asia, causing more than 10,000 hospitalizations annually, with a case-fatality rate 0.5-1.5% in Europe and 6-40% in Siberia and the Far East. A significant proportion of patients suffer from long-lasting neuropsychiatric sequelae. Inactivated vaccines produced in chick embryo cell cultures have proven effective in preventing the disease. For example, an 86% vaccination coverage of the Austrian population (the highest among European countries) has resulted in an approximately 90% reduction of hospitalized cases (Heinz and Kunz, *Arch. Virol. Suppl.* 18:201-205, 2004). The inactivated vaccines are expensive and require three inoculations for primary immunization. Periodic boosters (every 2-5 years) are required to maintain immunity. Therefore, a less

costly TBE vaccine, which is effective after one-two doses and provides durable, such as life-long immunity (similar to that achieved by YF 17D immunization) is needed, and indeed has been identified by the WHO as a major priority. Development of TBE LAV candidates in the past several decades by means of empirical or rational attenuation of TBE virus parent per se or chimerization of TBE or Langat (LGT, a naturally attenuated flavivirus that is closely related (serologically) to TBE) viruses with dengue 4 virus has faced difficulties due to problems with residual virulence of candidates and/or low immunogenicity/overattenuation (Wright et al., *Vaccine* 26:882-890, 2008; Maximova et al., *J. Virol.* 82:5255-5268, 2008; Rumyantsev et al., *Vaccine* 24:133-143, 2006; Kofler et al., *Arch. Virol. Suppl.* 18:191-200, 2004; and references therein).

Flaviviruses are small, enveloped, plus-strand RNA viruses transmitted primarily by arthropod vectors (mosquitoes or ticks) to natural hosts, which are primarily vertebrate animals, such as various mammals, including humans, and birds. The flavivirus genomic RNA molecule is about 11,000 nucleotides (nt) in length and encompasses a long open reading frame (ORF) flanked by 5' and 3' untranslated terminal regions (UTRs) of about 120 and 500 nucleotides in length, respectively. The ORF encodes a polyprotein precursor that is cleaved co- and post-translationally to generate individual viral proteins. The proteins are encoded in the order: C-prM/M-E-NS1-NS2A/2B-NS3-NS4A/4B-NS5, where C (core/capsid), prM/M (pre-membrane/membrane), and E (envelope) are the structural proteins, i.e., the components of viral particles, and the NS proteins are non-structural proteins, which are involved in intracellular virus replication. Flavivirus replication occurs in the cytoplasm. Upon infection of cells and translation of genomic RNA, processing of the polyprotein starts with translocation of the prM portion of the polyprotein into the lumen of endoplasmic reticulum (ER) of infected cells, followed by translocation of E and NS1 portions, as directed by the hydrophobic signals for the prM, E, and NS1 proteins. Amino-termini of prM, E, and NS1 proteins are generated by cleavage with cellular signalase, which is located on the luminal side of the ER membrane, and the resulting individual proteins remain carboxy-terminally anchored in the membrane. Most of the remaining cleavages, in the nonstructural region, are carried out by the viral NS2B/NS3 serine protease. The viral protease is also responsible for generating the C-terminus of the mature C protein found in progeny virions. Newly synthesized genomic RNA molecules and the C protein form a dense spherical nucleocapsid, which becomes surrounded by cellular membrane in which the E and prM proteins are embedded. The mature M protein is produced by cleavage of prM shortly prior to virus release by cellular furin or a similar protease. E, the major protein of the envelope, is the principal target for neutralizing antibodies, the main correlate of immunity against flavivirus infection. Virus-specific cytotoxic T-lymphocyte (CTL) response is the other key attribute of immunity. Multiple CD8+ and CD4+ CTL epitopes have been characterized in various flavivirus structural and non-structural proteins. In addition, innate immune responses contribute to both virus clearance and regulating the development of adaptive immune responses and immunologic memory.

In addition to the inactivated (INV) and live-attenuated (LAV) vaccines against flaviviruses discussed above, other vaccine platforms have been developed. One example is based on chimeric flaviviruses that include yellow fever virus capsid and non-structural sequences and prM-E proteins from other flaviviruses, to which immunity is sought. This technology has been used to develop vaccine candidates against

dengue (DEN), Japanese encephalitis (JE), West Nile (WN), and St. Louis encephalitis (SLE) viruses (see, e.g., U.S. Pat. Nos. 6,962,708 and 6,696,281). Yellow fever virus-based chimeric flaviviruses have yielded highly promising results in clinical trials.

Another flavivirus vaccine platform is based on the use of pseudoinfectious virus (PIV) technology (Mason et al., *Virology* 351:432-443, 2006; Shustov et al., *J. Virol.* 21:11737-11748, 2007; Widman et al., *Adv. Virus. Res.* 72:77-126, 2008; Suzuki et al., *J. Virol.* 82:6942-6951, 2008; Suzuki et al., *J. Virol.* 83:1870-1880, 2009; Ishikawa et al., *Vaccine* 26:2772-2781, 2008; Widman et al., *Vaccine* 26:2762-2771, 2008). PIVs are replication-defective viruses attenuated by a deletion(s). Unlike live flavivirus vaccines, they undergo a single round replication in vivo (or optionally limited rounds, for two-component constructs; see below), which may provide benefits with respect to safety. PIVs also do not induce viremia and systemic infection. Further, unlike inactivated vaccines, PIVs mimic whole virus infection, which can result in increased efficacy due to the induction of robust B- and T-cell responses, higher durability of immunity, and decreased dose requirements. Similar to whole viruses, PIV vaccines target antigen-presenting cells, such as dendritic cells, stimulate toll-like receptors (TLRs), and induce balanced Th1/Th2 immunity. In addition, PIV constructs have been shown to grow to high titers in substrate cells, with little or no cytopathic effect (CPE), allowing for high-yield manufacture, optionally employing multiple harvests and/or expansion of infected substrate cells.

The principles of the PIV technology are illustrated in FIGS. 1 and 2. There are two variations of the technology. In the first variation, a single-component pseudoinfectious virus (s-PIV) is constructed with a large deletion in the capsid protein (C), rendering mutant virus unable to form infectious viral particles in normal cells (FIG. 1). The deletion does not remove the first ~20 codons of the C protein, which contain an RNA cyclization sequence, and a similar number of codons at the end of C, which encode a viral protease cleavage site and the signal peptide for prM. The s-PIV can be propagated, e.g., during manufacture, in substrate (helper) cell cultures in which the C protein is supplied in trans, e.g., in stably transfected cells producing the C protein (or a larger helper cassette including C protein), or in cells containing an alphavirus replicon [e.g., a Venezuelan equine encephalitis virus (VEE) replicon] expressing the C protein or another intracellular expression vector expressing the C protein. Following inoculation in vivo, e.g., after immunization, the PIV undergoes a single round of replication in infected cells in the absence of trans-complementation of the deletion, without spread to surrounding cells. The infected cells produce empty virus-like particles (VLPs), which are the product of the prM-E genes in the PIV, resulting in the induction of neutralizing antibody response. A T-cell response should also be induced via MHC1 presentation of viral epitopes. This approach has been applied to YF 17D virus and WN viruses and WN/JE and WN/DEN2 chimeric viruses (Mason et al., *Virology* 351:432-443, 2006; Suzuki et al., *J. Virol.* 83:1870-1880, 2009; Ishikawa et al., *Vaccine* 26:2772-2781, 2008; Widman et al., *Vaccine* 26:2762-2771, 2008; WO 2007/098267; WO 2008/137163).

In the second variation, a two-component PIV (d-PIV) is constructed (FIG. 2). Substrate cells are transfected with two defective viral RNAs, one with a deletion in the C gene and another lacking the prM-E envelope protein genes. The two defective genomes complement each other, resulting in accumulation of two types of PIVs in the cell culture medium (Shustov et al., *J. Virol.* 21:11737-11748, 2007; Suzuki et al., *J. Virol.* 82:6942-6951, 2008). Optionally, the two PIVs can

be manufactured separately in appropriate helper cell lines and then mixed in a two-component formulation. The latter may offer an advantage of adjusting relative concentrations of the two components, increasing immunogenicity and efficacy. This type of PIV vaccine should be able to undergo a limited spread in vivo due to coinfection of some cells at the site of inoculation with both components. The spread is expected to be self-limiting as there are more cells in tissues than viral particles produced by initially coinfecting cells. In addition, a relatively high MOI is necessary for efficient coinfection, and cells outside of the inoculation site are not expected to be efficiently coinfecting (e.g., in draining lymph nodes). Cells infected with the AC PIV alone produce the highly immunogenic VLPs. Coinfecting cells produce the two types of packaged defective viral particles, which also stimulate neutralizing antibodies. The limited infection is expected to result in a stronger neutralizing antibody response and T-cell response compared to s-PIVs. To decrease chances of recombination during manufacture or in vivo, including with circulating flaviviruses, viral sequences can be modified in both s-PIVs and d-PIVs using, e.g., synonymous codon replacements, to reduce nucleotide sequence homologies, and mutating the complementary cyclization 5' and 3' elements.

SUMMARY OF THE INVENTION

The invention provides replication-deficient or defective pseudoinfectious flaviviruses including a flavivirus genome that includes (i) one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5), and (ii) sequences encoding one or more heterologous pathogen, cancer, or allergy-related immunogens. For example, the deletion/mutation can be within capsid (C) sequences; pre-membrane (prM) and/or envelope (E) sequences; capsid (C), pre-membrane (prM), and envelope (E) sequences; or non-structural protein 1 (NS1) sequences.

The heterologous immunogen can be, for example, from a pathogen selected from the group consisting of a rabies virus (e.g., a rabies virus G protein epitope), *Borrelia burgdorferi* (e.g., OspA immunogen or an immunogenic fragment thereof), a tick (e.g., a tick saliva protein selected from the group consisting of 64TRP, Isac, and Salp20, or an immunogenic fragment thereof), an influenza virus (e.g., an influenza virus M2, hemagglutinin (HA), or neuraminidase (NA) epitope, or an immunogenic fragment thereof), a human immunodeficiency virus (e.g., a codon-optimized HIV gag, pol, tat/nef, pro, or variants of Env protein, such as gp160, gp145, gp140, gp120, gp41, etc., or immunogenic fragments thereof), a simian immunodeficiency virus (e.g., a codon-optimized SIV gag, pol, tat/nef, pro, or variants of Env, or immunogenic fragments thereof), a human papilloma virus (e.g., an HPV16 or HPV18 capsid protein L1 or L2, or an immunogenic fragment thereof), a respiratory syncytial virus (e.g., a respiratory syncytial virus F or G glycoprotein), malaria parasite, and *Mycobacterium tuberculosis* (also see below).

The replication-deficient pseudoinfectious flaviviruses can include sequences encoding a pre-membrane (prM) and/or envelope (E) protein. Further, the replication-deficient pseudoinfectious flavivirus genomes can be selected from those of yellow fever virus, West Nile virus, tick-borne encephalitis virus, Langkat virus, Japanese encephalitis virus, dengue virus, and St. Louis encephalitis virus, attenuated

strains thereof, and chimeras thereof (also see below). In various examples, the chimeras include pre-membrane (prM) and envelope (E) sequences of a first flavivirus (e.g., a tick-borne encephalitis virus or a Langat virus), and capsid (C) and non-structural sequences of a second, different flavivirus (e.g., a yellow fever, a West Nile, or Langat virus).

The replication-deficient pseudoinfectious flavivirus genomes can be packaged in particles including pre-membrane (prM) and envelope (E) sequences from a flavivirus that is the same or different from that of the genomes. Further, the sequences encoding the heterologous immunogens can be inserted in the place of, or in combination with, the deletion(s) or mutation(s) of the one or more proteins.

The sequences encoding the heterologous immunogens can be inserted in the flavivirus genomes within sequences encoding the envelope (E) protein, within sequences encoding the non-structural 1 (NS1) protein, within sequences encoding the pre-membrane (prM) protein, intergenically between sequences encoding the envelope (E) protein and non-structural protein 1 (NS1), intergenically between non-structural protein 2B (NS2B) and non-structural protein 3 (NS3), and/or as a bicistronic insertion in the 3' untranslated region of the flavivirus genome.

In several embodiments, the replication-deficient pseudoinfectious flavivirus genomes include heterologous immunogen sequences from HIV, SIV, or influenza virus, such as any one or more of those described in Appendices 6-8. In particular embodiments, the replication-deficient pseudoinfectious virus is selected from any one of the SIV constructs 1-11 of Appendix 6, a construct having at least 50% sequence identity (e.g., 50%, 60%, 70%, 85%, 90%, 95%, or 99% or more sequence identity) to the nucleic acid or amino acid sequences described therein, or a construct that includes homologs and/or other naturally occurring variants of the SIV protein(s). In other embodiments, the replication-deficient pseudoinfectious virus is selected from the HIV Gag construct (PIV-WN (AprME)-HIV Gag) of Appendix 7, a construct having at least 50% sequence identity (e.g., 50%, 60%, 70%, 85%, 90%, 95%, or 99% or more sequence identity) to the nucleic acid or amino acid sequences described therein, or a construct that includes homologs and/or other naturally occurring variants of the HIV Gag protein. In still other embodiments, the replication-deficient pseudoinfectious virus is selected from the HIV Env construct (PIV-WN (AprME)-HIV Env Gp140) of Appendix 7, a construct having at least 50% sequence identity (e.g., 50%, 60%, 70%, 85%, 90%, 95%, or 99% or more sequence identity) to the nucleic acid or amino acid sequences described therein, or a construct that includes homologs and/or other naturally occurring variants of the HIV Env protein. In yet other embodiments, the replication-deficient pseudoinfectious virus is selected from construct 1 or 2 of Appendix 8, a construct having at least 50% sequence identity (e.g., 50%, 60%, 70%, 85%, 90%, 95%, or 99% or more sequence identity) to the nucleic acid or amino acid sequences described therein, or a construct that includes homologs and/or other naturally occurring variants of the HA protein.

The invention also includes compositions including a first replication-deficient pseudoinfectious flavivirus, as described above, and a second (or further), different replication-deficient pseudoinfectious flavivirus including a genome that includes one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5). In these compositions, the one or more proteins encoded by the

sequences in which the deletion(s) or mutation(s) occur in the second, different replication-deficient pseudoinfectious flavivirus are different from the one or more proteins encoded by the sequences in which the deletion(s) occur in the first replication-deficient pseudoinfectious flavivirus.

The invention further includes methods of inducing immune responses to an immunogen in a subject, which involves administering to the subject one or more replication-deficient pseudoinfectious flavivirus and/or composition as described herein to the subject. In particular embodiments, the replication-deficient pseudoinfectious flavivirus and/or composition includes any one or more of those described in Appendices 6-8, constructs having at least 50% sequence identity (e.g., 50%, 60%, 70%, 85%, 90%, 95%, or 99% or more sequence identity) to the nucleic acid or amino acid sequences described therein, or constructs that include homologs and/or other naturally occurring variants of the immunogenic SIV, HIV, and/or HA proteins. In various examples, the subject is at risk of but does not have an infection by the pathogen or a disease or condition associated with the cancer or allergy-related immunogen. In other examples, the subject has an infection by the pathogen or a disease or condition associated with the cancer or allergy-related immunogen. The invention thus includes prophylactic and therapeutic methods. In these methods, the immunogen can be from, for example, a pathogen selected from the group consisting of a rabies virus, *Borrelia burgdorferi*, a tick, an influenza virus, a human immunodeficiency virus, a simian immunodeficiency virus, a human papilloma virus, a respiratory syncytial virus, malaria parasite, and *Mycobacterium tuberculosis* (also see below). Further, the methods can be for inducing an immune response against a protein encoded by the flavivirus genome, in addition to the source of the immunogen. In various examples, the subject is at risk of but does not have an infection by the flavivirus corresponding to the genome of the pseudoinfectious flavivirus, which includes sequences encoding a flavivirus pre-membrane and/or envelope protein. In other examples, the subject has an infection by the flavivirus corresponding to the genome of the pseudoinfectious flavivirus, which includes sequences encoding a flavivirus pre-membrane and/or envelope protein.

The invention also includes live, attenuated chimeric flaviviruses including a yellow fever virus in which sequences encoding pre-membrane and envelope proteins are replaced with sequences encoding pre-membrane and envelope proteins of a tick-borne encephalitis virus or a Langat virus, and the signal sequence between the capsid and pre-membrane proteins of the chimeric flavivirus includes a hybrid of yellow fever virus and tick-borne encephalitis or Langat virus capsid/pre-membrane signal sequences, or a variant thereof. In various examples, the capsid/pre-membrane signal sequence of the chimeric flavivirus includes yellow fever virus sequences in the amino terminal region and tick-borne encephalitis or Langat virus sequences in the carboxy terminal region (see below).

Further, the invention includes live, attenuated chimeric flaviviruses including a West Nile virus in which sequences encoding pre-membrane and envelope proteins are replaced with sequences encoding pre-membrane and envelope proteins of a tick-borne encephalitis or a Langat virus, and the signal sequence between the capsid and pre-membrane proteins of the chimeric flavivirus includes a tick-borne encephalitis or a Langat virus capsid/pre-membrane signal sequence, or a variant thereof.

The invention also includes pharmaceutical compositions including one or more pseudoinfectious flavivirus, composition, or live, attenuated flavivirus as described herein, and a

pharmaceutically acceptable carrier or diluent. Further, the compositions can include an adjuvant.

Also included in the invention are replication-deficient pseudoinfectious flaviviruses including a flavivirus genome including one or more deletion(s) or mutation(s) in nucleotide sequences encoding non-structural protein 1 (NS1), non-structural protein 3 (NS3), or non-structural protein 5 (NS5).

Further, the invention includes nucleic acid molecules corresponding to the genome of a pseudoinfectious flavivirus, or the genome of the live, attenuated flavivirus, as described herein, and complements thereof.

The invention also provides methods of making replication-deficient pseudoinfectious flaviviruses as described herein, involving introducing one or more nucleic acid molecules, as described above, into a cell that expresses the protein(s) corresponding to any sequences deleted from the flavivirus genome of the replication-deficient pseudoinfectious flaviviruses. In these methods, the protein can be expressed in the cell from the genome of a second (or further), different, replication-deficient pseudoinfectious flavivirus. In other examples, the protein is expressed from a replicon (e.g., an alphavirus replicon, such as a Venezuelan Equine Encephalitis virus replicon; see below).

The invention also includes compositions containing two or more replication-deficient pseudoinfectious flaviviruses, in which two of the replication-deficient pseudoinfectious flaviviruses are selected from the groups consisting of: (a) a replication-deficient pseudoinfectious flavivirus including a genome containing Japanese encephalitis virus sequences, and a replication-deficient pseudoinfectious flavivirus including a genome containing dengue virus sequences; (b) a replication-deficient pseudoinfectious flavivirus including a genome containing yellow fever virus sequences, and a replication-deficient pseudoinfectious flavivirus including a genome containing dengue virus sequences; and (c) a replication-deficient pseudoinfectious flavivirus including a genome containing tick-borne encephalitis or Langat virus sequences and an inserted sequence encoding a *Borrelia burgdorferi* immunogen, and a replication-deficient pseudoinfectious flavivirus including a genome containing tick-borne encephalitis or Langat virus sequences and an inserted sequence encoding a tick saliva protein immunogen, or a replication-deficient pseudoinfectious flavivirus including a genome containing tick-borne encephalitis or Langat virus sequences and inserted sequences encoding a *Borrelia burgdorferi* immunogen and a tick saliva protein immunogen.

Pharmaceutical compositions including the live, attenuated chimeric flaviviruses described herein are also included in the invention. Further, the invention includes methods of inducing an immune response to tick-borne encephalitis virus or Langat virus in a subject, involving administering to the subject such a pharmaceutical composition. In various examples, the subject does not have but is at risk of developing infection by tick-borne encephalitis virus or Langat virus. In other examples, the subject is infected with tick-borne encephalitis virus or Langat virus.

The invention further includes replication-deficient pseudoinfectious flaviviruses including a flavivirus genome including one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5), wherein the flavivirus genome includes yellow fever virus sequences in which sequences encoding pre-membrane and envelope proteins are replaced with sequences encoding pre-membrane and envelope proteins of a tick-borne encephalitis virus or a

Langat virus, and sequences encoding the signal sequence between the capsid and pre-membrane proteins of the flavivirus genome include a hybrid of sequences encoding yellow fever virus and tick-borne encephalitis or Langat virus capsid/pre-membrane signal sequences, or a variant thereof. In various examples, the sequences encoding the capsid/pre-membrane signal sequence of the flavivirus genome include yellow fever virus sequences in the 5' region and tick-borne encephalitis or Langat virus sequences in the 3' region.

Further, the invention includes replication-deficient pseudoinfectious flaviviruses including a flavivirus genome including one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5), wherein the flavivirus genome includes West Nile virus sequences in which sequences encoding pre-membrane and envelope proteins are replaced with sequences encoding pre-membrane and envelope proteins of a tick-borne encephalitis or a Langat virus, and the sequences encoding the signal sequence between the capsid and pre-membrane proteins of the flavivirus genome include sequences encoding a tick-borne encephalitis or a Langat virus capsid/pre-membrane signal sequence, or a variant thereof.

In addition, the invention includes replication-deficient pseudoinfectious flaviviruses including a flavivirus genome including one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5), wherein any capsid (C) and non-structural (NS) proteins in the flavivirus genome are from Langat virus and any pre-membrane (prM) and envelope (E) proteins are from a tick-borne encephalitis virus.

By "replication-deficient pseudoinfectious flavivirus" or "PIV" is meant a flavivirus that is replication-deficient due to a deletion or mutation in the flavivirus genome. The deletion or mutation can be, for example, a deletion of a large sequence, such as most of the capsid protein, as described herein (with the cyclization sequence remaining; see below). In other examples, sequences encoding different proteins (e.g., prM, E, NS1, NS3, and/or NS5; see below) or combinations of proteins (e.g., prM-E or C-prM-E) are deleted. This type of deletion may be advantageous if the PIV is to be used a vector to deliver a heterologous immunogen, as the deletion can permit insertion of sequences that may be, for example, at least up to the size of the deleted sequence. In other examples, the mutation can be, for example, a point mutation, provided that it results in replication deficiency, as discussed above. Because of the deletion or mutation, the genome does not encode all proteins necessary to produce a full flavivirus particle. The missing sequences can be provided in trans by a complementing cell line that is engineered to express the missing sequence (e.g., by use of a replicon; s-PIV; see below), or by co-expression of two replication-deficient genomes in the same cell, where the two replication-deficient genomes, when considered together, encode all proteins necessary for production (d-PIV system; see below).

Upon introduction into cells that do not express complementing proteins, the genomes replicate and, in some instances, generate "virus-like particles," which are released from the cells and are able to leave the cells and be immunogenic, but cannot infect other cells and lead to the generation of further particles. For example, in the case of a PIV including a deletion in capsid protein encoding sequences, after

infection of cells that do not express capsid, VLPs including prM-E proteins are released from the cells. Because of the lack of capsid protein, the VLPs lack capsid and a nucleic acid genome. In the case of the d-PIV approach, production of further PIVs is possible in cells that are infected with two PIVs that complement each other with respect to the production of all required proteins (see below).

Also included in the invention are replication-defective pseudoinfectious flaviviruses including multiple heterologous immunogens from, e.g., a human immunodeficiency virus or a simian immunodeficiency virus. In various examples, the multiple immunogens can include heterologous transmembrane and/or signal sequences (from, e.g., a rabies virus G protein).

The invention provides several advantages. For example, the PIV vectors and PIVs of the invention are highly attenuated and highly efficacious after one-to-two doses, providing durable immunity. Further, unlike inactivated vaccines, PIVs mimic whole virus infection, which can result in increased efficacy due to the induction of robust B- and T-cell responses, higher durability of immunity, and decreased dose requirements. In addition, similar to whole viruses, PIV vaccines target antigen-presenting cells, such as dendritic cells, stimulate toll-like receptors (TLRs), and induce balanced Th1/Th2 immunity. PIV constructs have also been shown to grow to high titers in substrate cells, with little or no CPE, allowing for high-yield manufacture, optionally employing multiple harvests and/or expansion of infected substrate cells. Further, the PIV vectors of the invention provide an option for developing vaccines against non-flavivirus pathogens for which no vaccines are currently available.

Other features and advantages of the invention will be apparent from the following detailed description, the drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic illustration of single component PIV (s-PIV) technology.

FIG. 2 is a schematic illustration of two-component PIV (d-PIV) technology.

FIG. 3 is a schematic illustration of a general experimental design for testing immunogenicity and efficacy of PIVs in mice.

FIG. 4 is a graph comparing the humoral immune response induced by PIV-WN (RV-WN) with that of YF/WN LAV (CV-WN) in mice.

FIG. 5 is a series of graphs showing the results of challenging hamsters immunized with PIV-YF (RV-YF), YF17D, PIV-WN(RV-WN), and YF/WN LAV (CVWN) with hamster-adapted Asibi (PIV-YF and YF 17D vaccinees) and wild type WN-NY99 (PIV-WN and YF/WN LAV vaccinees).

FIG. 6 is a table showing YF/TBE and YF/LGT virus titers and plaque morphology obtained with the indicated chimeric flaviviruses.

FIG. 7 is a table showing WN/TBE PIV titers and examples of immunofluorescence of cells containing the indicated PIVs.

FIG. 8 is a set of graphs showing the replication kinetics of YF/TBE LAV and PIV-WN/TBE in Vero and BHK cell lines (CV-Hypr=YF/Hypr LAV; CV-LGT=YF/LGT LAV; RV-WN/TBEV=PIV-WN/TBEV).

FIG. 9 is a series of graphs showing survival of mice inoculated IC with PIV-TBE and YF/TBE LAV constructs in a neurovirulence test (3.5 week old ICR mice; RV-WN/Hypr=PIV-WN/TBE(Hypr); CV-Hypr=YF/TBE(Hypr) LAV; CV-LGT=YF/LGT LAV).

FIG. 10 is a graph showing survival of mice inoculated IP with PIV-WN/TBE(Hypr) (RV-WN/Hypr), YF/TBE(Hypr) LAV (CV-Hypr), and YF/LGT LAV (CV-LGT) constructs and YF17D in a neuroinvasiveness test (3.5 week old ICR mice).

FIG. 11 is a series of graphs illustrating morbidity in mice measured by dynamics of body weight loss after TBE virus challenge, for groups immunized with S-PIV-TBE candidates (upper left panel), YF/TBE and YF/LGT chimeric viruses (upper right panel), and controls (YF 17D, human killed TBE vaccine, and mock; bottom panel).

FIG. 12 is a schematic representation of PIV constructs expressing rabies virus G protein, as well as illustration of packaging of the constructs to make pseudoinfectious virus and immunization.

FIG. 13 is a schematic representation of insertion designs resulting in viable/expressing constructs (exemplified by rabies G).

FIG. 14 is series of images showing immunofluorescence analysis and graphs showing growth curves of cells transfected with the indicated PIV-WN constructs (Δ C-Rabies G, Δ PrM-E-Rabies G, and Δ C-PrM-E-Rabies G).

FIG. 15 is a series of images showing immunofluorescence analysis of RabG expressed on the plasma membranes of Vero cells transfected with the indicated PIV constructs (Δ C-Rabies G, Δ PrM-E-Rabies G, and AC-PrM-E-Rabies G).

FIG. 16 is a schematic illustration of a PIV-WN-rabies G construct and a series of images showing that this construct spreads in helper cells, but not in naïve cells.

FIG. 17 is a series of graphs showing stability of the rabies G protein gene in PIV-WN vectors.

FIG. 18 is a set of images showing a comparison of spread of single-component vs. two-component PIV-WN-rabies G variants in Vero cells.

FIG. 19 is a set of immunofluorescence images showing expression of full-length RSV F protein (strain A2) by the AprM-E component of d-PIV-WN in helper cells after transfection.

FIG. 20 is a schematic representation of an artificial cassette containing SIV (GenBank accession number ADM52218.1) gp120 (the native signal sequence in the gene was replaced with the tPA signal and gp41 was truncated to contain only the TM domain), Gag, and Pro (protease) genes.

FIG. 21 is a schematic representation of inserts of the first three constructs in FIG. 20 (the three top constructs shown in FIG. 21), starting with the Env glycoprotein that were designed similarly to the PIV WN-rabies G vectors described herein (see, e.g., FIGS. 12-14 and hereinbelow), in which the gp120 signal is fused with a portion of the signal sequence for prM (e.g., at the end of the C gene or downstream from Δ C deletion depending on vector). In addition, schematic representations of alternate dC RV230 Env PIV constructs are shown (the three bottom constructs shown in FIG. 21).

FIG. 22 is a schematic representation of Gag and Gag-Pro PIV construct designs, in which Gag and Gag-Pro were cloned in place of the Δ prM-E or Δ C-prM-E deletions.

FIG. 23 is a photograph of a Western blot using anti-Gag antibodies, which shows correct processing of the polyprotein in recovered SIV Gag and SIV Gag/Pro PIVs grown in helper cells.

FIGS. 24A-24C are photomicrographs showing that immunostaining of naïve Vero cells infected with the Gag PIVs, showed individual stained cells as expected from sPIV. FIG. 24A is a negative control, FIG. 24B shows immunostaining of naïve Vero cells infected with RV230 9AA-FMD-Gag PIV, and FIG. 24C shows immunostaining of

naïve Vero cells infected with RV230 FMD-Gag PIV. The two constructs are illustrated schematically in FIG. 24D.

FIGS. 25A-F are graphs showing growth curves of SIV Gag PIV variants after transfection of helper cells with in vitro synthesized PIV RNA (P0 passage) indicating efficient replication in vivo. Immunofluorescence images of Vero cells infected with the variants are shown inset.

FIG. 26 is a graph showing growth curves in naïve Vero cells of SIV Gag PIV as a two-component formulation (d-PIV, sometimes also designated as tc-PIV) together with PIV-WN helper with ΔC deletion (RV909).

FIG. 27 is a graph showing high insert stability for one of the SIV Gag PIV variants (RV230-Gag variant, containing Gag gene in place of large $\Delta prM-E$ deletion, in helper BHK-CprME(WN) cells at MOI 0.1 FFU/cell) when examined by ten serial passages.

FIGS. 28A-D are immunofluorescence images showing efficient expression of SIV Env (gp120) in Vero cells using PIV-(WN)-SIV Env variants. Efficient intracellular expression of the original gp120 was observed in Vero cells infected with packaged dC230Env PIV variant as determined by immunostaining using anti-SIV Env antibody after methanol fixation (FIG. 28D), although transport of gp120 to the surface of infected Vero cells was inefficient, as determined following formalin fixation (FIG. 28B). In contrast, the dC230Env/RabG anchor PIV construct (see FIG. 21), in which the SIV Env TM domain was replaced with the TM anchor sequence from rabies virus G protein, showed efficient intracellular (FIG. 28C) and extracellular expression (FIG. 28A).

FIG. 29 is an immunofluorescence image showing expression of SIV Env on the surface of PIV-SIV Env/RabG TM infected Vero cells.

FIG. 30 is a schematic representation of PIV-flu HA construct designs, in which the full-length HA gene of Flu strain New Calcdonia was cloned in place of $\Delta prM-E$ and $\Delta C-prM-E$ deletions of PIV-WN vectors in the same fashion as described for Rabies G, RSV F and SIV Env (as is described herein).

FIGS. 31A-B are graphs showing growth curves in BHK 363 helper cells transfected at P14 with RNA from RV230 HA New Calcdonia PIV clones 6 (FIG. 31A) and 10 (FIG. 31B), as determined by immunostaining with anti-WN and anti-HA antibodies.

FIGS. 32A-D are graphs showing growth curves in BHK 363 helper cells transfected at P14 with RNA from RV230 HA New Calcdonia PIV clones 1, 6, and 10 (FIGS. 32A-C, respectively) and from dC RV230 HA New Calcdonia PIV clone 6 (FIG. 32D), as determined by immunostaining with anti-WN and anti-HA antibodies.

FIGS. 33A-F are immunofluorescence images showing surface expression (FIGS. 33A-C) and intracellular expression (FIGS. 33D-F) of HA in Vero cells infected with RV230 HA New Calcdonia PIV clones 1, 6, and 10, respectively.

FIGS. 34A-B are immunofluorescence images showing surface expression (FIG. 34A) and intracellular expression (FIG. 34B) of HA in Vero cells infected with dC RV230 HA New Calcdonia PIV clone 6.

FIG. 35 shows immunofluorescence images confirming surface expression (FIGS. 35B and D) and intracellular expression (FIGS. 35F and H) of HA in Vero cells infected with RV230 HA New Calcdonia PIV. FIGS. 35A, C, E, and G are negative controls showing the lack surface expression (FIGS. 35A and C) and intracellular expression (FIGS. 35E and G) of HA in uninfected Vero cells. The immunofluorescence images in FIGS. 35B and F were produced using antibodies against the stem of HA, while the immunofluores-

cence images in FIGS. 35D and H were produced using antibodies against the HA globular head. FIGS. 35B, D, F, and H confirm the correct, native protein confirmation of HA.

FIGS. 36A-D are immunofluorescence images showing surface expression (FIGS. 36A and B) and intracellular expression (FIGS. 36C and D) of HA in Vero cells infected with RV230 HA New Calcdonia PIV clones 6 and 10, respectively, 48 hours post infection. Staining was performed with a mix of HA stem and globular head antibodies.

FIG. 37A is an immunofluorescence image showing staining of RV230-HA PIV infected Vero cells by HA stem-specific antibodies. FIG. 37B is an immunofluorescence image showing staining of RV230-HA PIV infected Vero cells by HA globular head-specific antibodies.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides replication-defective or deficient pseudoinfectious virus (PIV) vectors including flavivirus sequences, which can be used in methods for inducing immunity against heterologous pathogen, cancer, and allergy-related immunogens inserted into the vectors as well as, optionally, the vectors themselves. The invention also includes compositions including combinations of PIVs and/or PIV vectors, as described herein, and methods of using such compositions to induce immune responses against inserted immunogen sequences and/or sequences of the PIVs themselves. Further, the invention includes particular PIVs and live, attenuated chimeric flaviviruses including tick-borne encephalitis virus sequences, and related vectors, compositions, and methods of use. The PIV vectors, PIVs, live attenuated chimeric flaviviruses, compositions, and methods of the invention are described further below.

PIV Vectors and PIVs

The PIV vectors and PIVs of the invention can be based on the single- or two-component PIVs described above (also see WO 2007/098267 and WO 2008/137163). Thus, for example, in the case of single component PIVs, the PIV vectors and PIVs can include a genome including a large deletion in capsid protein encoding sequences and be produced in a complementing cell line that produces capsid protein in trans (single component; FIG. 1 and FIG. 12). According to this approach, most of the capsid-encoding region is deleted, which prevents the PIV genome from producing infectious progeny in normal cell lines (i.e., cell lines not expressing capsid sequences) and vaccinated subjects. The capsid deletion typically does not disrupt RNA sequences required for genome cyclization (i.e., the sequence encoding amino acids in the region of positions 1-26), and/or the prM sequence required for maturation of prM to M. In specific examples, the deleted sequences correspond to those encoding amino acids 26-100, 26-93, 31-100, or 31-93 of the C protein.

Single component PIV vectors and PIVs can be propagated in cell lines that express either C or a C-prM-E cassette, where they replicate to high levels. Exemplary cell lines that can be used for expression of single component PIV vectors and PIVs include BHK-21 (e.g., ATCC CCL-10), Vero (e.g., ATCC CCL-81), C7/10, and other cells of vertebrate or mosquito origin. The C or C-prM-E cassette can be expressed in such cells by use of a viral vector-derived replicon, such as an alphavirus replicon (e.g., a replicon based on Venezuelan Equine Encephalitis virus (VEEV), Sindbis virus, Semliki Forest virus (SFV), Eastern Equine Encephalitis virus (EEEV), Western Equine Encephalitis virus (WEEV), or Ross River virus). To decrease the possibility of productive recombination between the PIV vectors/PIVs and complementing sequences, the sequences in the replicons (encoding

C, prM, and/or E) can include nucleotide mutations. For example, sequences encoding a complementing C protein can include an unnatural cyclization sequence. The mutations can result from codon optimization, which can provide an additional benefit with respect to PIV yield. Further, in the case of complementing cells expressing C protein sequences (and not a C-prM-E cassette), it may be beneficial to include an anchoring sequence at the carboxy terminus of the C protein including, for example, about 20 amino acids of prM (see, e.g., WO 2007/098267).

The PIV vectors and PIVs of the invention can also be based on the two-component genome technology described above. This technology employs two partial genome constructs, each of which is deficient in expression of at least one protein required for productive replication (capsid or prM/E) but, when present in the same cell, result in the production of all components necessary to make a PIV. Thus, in one example of the two-component genome technology, the first component includes a large deletion of C, as described above in reference to single component PIVs, and the second component includes a deletion of prM and E (FIG. 2 and FIG. 12). In another example, the first component includes a deletion of C, prM, and E, and the second component includes a deletion of NS1 (FIG. 12). Both components can include cis-acting promoter elements required for RNA replication and a complete set of non-structural proteins, which form the replicative enzyme complex. Thus, both defective genomes can include a 5'-untranslated region and at least about 60 nucleotides (Element 1) of the following, natural protein-coding sequence, which comprises an amino-terminal fragment of the capsid protein. This sequence can be followed by a protease cleavage sequence such as, for example, a ubiquitin or foot-and-mouth disease virus (FAMDV)-specific 2A protease sequence, which can be fused with either capsid or envelope (prM-E) coding sequences. Further, artificial, codon optimized sequences can be used to exclude the possibility of recombination between the two defective viral genomes, which could lead to formation of replication-competent viruses (see, e.g., WO 2008/137163). Use of the two-component genome approach does not require the development of cell lines expressing complementing genomes, such as the cells transformed with replicons, as discussed above in reference to the single component PIV approach. Exemplary cell lines that can be used in the two-component genome approach include Vero (e.g., ATCC CCL-81), BHK-21 (e.g., ATCC CCL-10), C7/10, and other cells of vertebrate or mosquito origin.

Additional examples of d-PIV approaches that can be used in the invention are based on use of complementing genomes including deletions in NS3 or NS5 sequences. A deletion in, e.g., NS1, NS3, or NS5 proteins can be used as long as several hundred amino acids in the ORF, removing the entire chosen protein sequence, or as short as 1 amino acid inactivating protein enzymatic activity (e.g., NS5 RNA polymerase activity, NS3 helicase activity, etc.). Alternatively, point amino acid changes (as few as 1 amino acid mutation, or optionally more mutations) can be introduced into any NS protein, inactivating enzymatic activity. In addition, several Δ NS deletions can be combined in one helper molecule. The same heterologous gene, i.e., expressed by the first d-PIV component, can be expressed in place or in combination with the NS deletion(s) in the second component, increasing the amount of expressed immunogen. Notably, the insertion capacity of the helper will increase proportionally to the size of NS deletion(s). Alternatively, a different foreign immunogen(s) can be inserted in place of deletion(s) of the helper to produce multivalent vaccines.

Further, additional approaches that can be used in making PIV vectors and PIVs for use in the present invention are described, for example, in WO 99/28487, WO 03/046189, WO 2004/108936, US 2004/0265338, US 2007/0249032, and U.S. Pat. No. 7,332,322.

The PIV vectors and PIVs of the invention can be comprised of sequences from a single flavivirus type (e.g., tick-borne encephalitis (TBE, e.g., strain Hypr), Langat (LGT), yellow fever (e.g., YF17D), West Nile, Japanese encephalitis, dengue (serotype 1-4), St. Louis encephalitis, Kunjin, Rocio encephalitis, Ilheus, Central European encephalitis, Siberian encephalitis, Russian Spring-Summer encephalitis, Kyasanur Forest Disease, Omsk Hemorrhagic fever, Louping ill, Powassan, Negishi, Absettarov, Hansalova, and Apoi viruses), or can comprise sequences from two or more different flaviviruses. Sequences of some strains of these viruses are readily available from generally accessible sequence databases; sequences of other strains can be easily determined by methods well known in the art. In the case of PIV vectors and PIVs including sequences of more than one flavivirus, the sequences can be those of a chimeric flavivirus, as described above (also see, e.g., U.S. Pat. Nos. 6,962,708; 6,696,281; and 6,184,024). In certain examples, the chimeras include pre-membrane and envelope sequences from one flavivirus (such as a flavivirus to which immunity may be desired), and capsid and non-structural sequences from a second, different flavivirus. In one specific example, the second flavivirus is a yellow fever virus, such as the vaccine strain YF17D. Other examples include the YF/TBE, YF/LGT, WN/TBE, and WN/LGT chimeras described below. Another example is an LGT/TBE chimera based on LGT virus backbone containing TBE virus prM-E proteins. A PIV vaccine based on this genetic background would have an advantage, because LGT replicates very efficiently in vitro and is highly attenuated and immunogenic for humans. Thus, a chimeric LGT/TBE PIV vaccine is expected to provide a robust specific immune response in humans against TBE, particularly due to inclusion of TBE prM-E genes. Vectors of the invention can be based on PIV constructs or live, attenuated chimeric flaviviruses as described herein (in particular, YF/TBE, YF/LGT, WN/TBE, and WN/LGT; see below). Use of PIV constructs as vectors provides particular advantages in certain circumstances, because these constructs by necessity include large deletions, which render the constructs more amenable to accommodation of insertions that are at least up to the size of the deleted sequences, without there being a loss in replication efficiency. Thus, PIV vectors in general can comprise very small insertions (e.g., in the range 6-10, 11-20, 21-100, 101-500, or more amino acid residues combined with the Δ C deletion or other deletions), as well as relatively large insertions or insertions of intermediate size (e.g., in the range 501-1000, 1001-1700, 1701-3000, or 3001-4000 or more residues). In contrast, in certain examples, it may be advantageous to express relatively short sequences in live attenuated viruses, particularly if the insertions are made in the absence of a corresponding deletion. Additional information concerning insertion sites that can be used in the invention is provided below. In addition, as discussed further below, expression of non-flavivirus immunogens in PIVs and chimeric flaviviruses of the invention can result in dual vaccines that elicit protective immunity against both a flavivirus vector virus pathogen and a target heterologous immunogen (e.g., a pathogen (such as a bacterial, viral, parasite, or fungal pathogen), cancer, or allergy-related immunogen).

As discussed above, the PIV vectors and PIVs of the invention can comprise sequences of chimeric flaviviruses, for example, chimeric flaviviruses including pre-membrane and

envelope sequences of a first flavivirus (e.g., a flavivirus to which immunity is sought), and capsid and non-structural sequences of a second, different flavivirus, such as a yellow fever virus (e.g., YF17D; see above and also U.S. Pat. No. 6,962,708; 6,696,281; and 6,184,024). Further, chimeric flaviviruses of the invention, used as a source for constructing PIVs, or as vaccines/vaccine vectors per se, can optionally include one or more specific attenuating mutations (e.g., E protein mutations, prM protein mutations, deletions in the C protein, and/or deletions in the 3'UTR), such as any of those described in WO 2006/116182. For example, the C protein or 3'UTR deletions can be directly applied to YF/TBE or YF/LGT chimeras. Similar deletions can be designed and introduced in other chimeric LAV candidates such as based on LGT/TBE, WN/TBE, and WN/LGT genomes. With respect to E protein mutations, attenuating mutations similar to those described for YF/WN chimera in WO 2006/116182 can be designed, e.g., based on the knowledge of crystal structure of the E protein (Rey et al., *Nature* 375(6529):291-298, 1995), and employed. Further, additional examples of attenuating E protein mutations described for TBE virus and other flaviviruses are provided in Table 9. These can be similarly introduced into chimeric vaccine candidates.

The invention also provides new, particular chimeric flaviviruses, which can be used as a basis for the design of PIV vectors and PIVs, as live attenuated chimeric flavivirus vectors, and as vaccines against the source(s) of the pre-membrane and envelope components of the chimeras. These chimeras include tick-borne encephalitis (TBE) virus or related prM-E sequences. Thus, the chimeras can include prM-E sequences from, for example, the Hypr strain of TBE or Langat (LGT) virus. Capsid and non-structural proteins of the chimeras can include those from yellow fever virus (e.g., YF17D) or West Nile virus (e.g., NY99).

A central feature of these exemplary YF/TBE, YF/LGT, WN/TBE, and WN/LGT chimeras is the signal sequence between the capsid and prM proteins. As is shown in the Examples, below, we have found that, in the case of YF-based PIV chimeras, it is advantageous to use a signal sequence comprising yellow fever and TBE sequences (see below). In one example, the signal sequence includes yellow fever sequences in the amino terminal region (e.g., SHDVLTVQ-FLIL) and TBE sequences in the carboxy terminal region (e.g., GMLGMTIA), resulting in the sequence SHDVLTVQ-FLILGMLGMTIA. We have also found that, in the case of WN-based PIV chimeras, it is advantageous to use a signal sequence comprising TBE sequences (e.g., GGTDWMSWLLVIGMLGMTIA). The invention thus includes YF/TBE, YF/LGT, WN/TBE, and WN/LGT chimeras, both PIVs and LAVs, which include the above-noted signal sequences, or variants thereof having, e.g., 1-8, 2-7, 3-6, or 4-5 amino acid substitutions, deletions, or insertions, which do not substantially interfere with processing at the signal sequence. In various examples, the substitutions are "conservative substitutions," which are characterized by replacement of one amino acid residue with another, biologically similar residue. Examples of conservative substitutions include the substitution of one hydrophobic residue such as isoleucine, valine, leucine, or methionine for another, or the substitution of one polar residue for another, such as between arginine and lysine, between glutamic and aspartic acids, or between glutamine and asparagine and the like. Examples of exemplary PIVs of the present invention include those described in Appendices 6-8, constructs having at least 50% sequence identity (e.g., 50%, 60%, 70%, 85%, 90%, 95%, or 99% or more sequence identity) to the nucleic acid or amino acid sequences described therein, or constructs that include

homologs and/or other naturally occurring variants of the SIV, HIV, and/or HA proteins. Additional information concerning these and other chimeras is provided below, in the Examples.

5 Insertion Sites

Sequences encoding immunogens can be inserted at one or more different sites within the vectors of the invention. Relatively short peptides can be delivered on the surface of PIV or LAV glycoproteins (e.g., prM, E, and/or NS1 proteins) and/or in the context of other proteins (to induce predominantly B-cell and T-cell responses, respectively). Other inserts, including larger portions of foreign proteins, as well as complete proteins, can be expressed intergenically, at the N- and C-termini of the polyprotein, or bicistronically (e.g., within the ORF under an IRES or in the 3'UTR under an IRES; see, e.g., WO 02/102828, WO 2008/036146, WO 2008/094674, WO 2008/100464, WO 2008/115314, and below for further details). In PIV constructs, there is an additional option of inserting a foreign amino acid sequence directly in place of introduced deletion(s). Insertions can be made in, for example, AC, AprM-E, AC-prM-E, ΔNS1, ΔNS3, and ΔNS5. Thus, in one example, in the case of s-PIVs and the AC component of d-PIVs, immunogen-encoding sequences can be inserted in place of deleted capsid sequences. Immunogen-encoding sequences can also, optionally, be inserted in place of deleted prM-E sequences in the AprM-E component of d-PIVs. In another example, the sequences are inserted in place of or combined with deleted sequences in ΔC-prM-E constructs. Examples of such insertions are provided in the Examples section, below.

In the case of making insertions into PIV deletions, the insertions can be made with a few (e.g., 1, 2, 3, 4, or 5) additional vector-specific residues at the N- and/or C-termini of the foreign immunogen, if the sequence is simply fused in-frame (e.g., ~20 first a.a. and a few last residues of the C protein if the sequence replaces the ΔC deletion), or without, if the foreign immunogen is flanked by appropriate elements well known in the field (e.g., viral protease cleavage sites; cellular protease cleavage sites, such as signalase, furin, etc.; autoprotease; termination codon; and/or IRES elements).

If a protein is expressed outside of the continuous viral open reading frame (ORF), e.g., if vector and non-vector sequences are separated by an internal ribosome entry site (IRES), cytoplasmic expression of the product can be achieved or the product can be directed towards the secretory pathway by using appropriate signal/anchor segments, as desired. If the protein is expressed within the vector ORF, important considerations include cleavage of the foreign protein from the nascent polyprotein sequence, and maintaining correct topology of the foreign protein and all viral proteins (to ensure vector viability) relative to the ER membrane, e.g., translocation of secreted proteins into the ER lumen, or keeping cytoplasmic proteins or membrane-associated proteins in the cytoplasm/in association with the ER membrane.

In more detail, the above-described approaches to making insertions can employ the use of, for instance, appropriate vector-derived, insert-derived, or unrelated signal and anchor sequences included at the N and C termini of glycoprotein inserts. For example, all or a portion of the rabies G-derived signal and/or anchor sequences can be used in place of all or a portion of the signal and/or anchor sequences for glycoprotein inserts (e.g., one or more of the SIV, HIV, or influenza virus proteins described herein) to produce a heterologous polypeptide sequence. Standard autoproteases, such as FMDV 2A autoprotease (~20 amino acids) or ubiquitin (gene

~500 nt), or flanking viral NS2B/NS3 protease cleavage sites can be used to direct cleavage of an expressed product from a growing polypeptide chain, to release a foreign protein from a vector polypeptide, and to ensure viability of the construct. Optionally, growth of the polypeptide chain can be terminated by using a termination codon, e.g., following a foreign gene insert, and synthesis of the remaining proteins in the constructs can be re-initiated by incorporation of an IRES element, e.g., the encephalomyocarditis virus (EMCV) IRES commonly used in the field of RNA virus vectors. Viable recombinants can be recovered from helper cells (or regular cells for d-PIV versions). Optionally, backbone PIV sequences can be rearranged, e.g., if the latter results in more efficient expression of a foreign gene. For example, a gene rearrangement has been applied to TBE virus, in which the prM-E genes were moved to the 3' end of the genome under the control of an IRES (Orlinger et al., J. Virol. 80:12197-12208, 2006). Translocation of prM-E or any other genes can be applied to PIV flavivirus vaccine candidates and expression vectors, according to the invention.

Additional details concerning different insertion sites that can be used in the invention are as follows (also see WO 02/102828, WO 2008/036146, WO 2008/094674, WO 2008/100464, WO 2008/115314, as noted above). Peptide sequences can be inserted within the envelope protein, which is the principle target for neutralizing antibodies. The sequences can be inserted into the envelope in, for example, positions corresponding to amino acid positions 59, 207, 231, 277, 287, 340, and/or 436 of the Japanese encephalitis virus envelope protein (see, e.g., WO 2008/115314 and WO 02/102828). To identify the corresponding loci in different flaviviruses, the flavivirus sequences are aligned with that of Japanese encephalitis virus. As there may not be an exact match, it should be understood that, in non-JE viruses, the site of insertion may vary by, for example, 1, 2, 3, 4, or 5 amino acids, in either direction. Further, given the identification of such sites as being permissive in JE, they can also vary in JE by, for example, 1, 2, 3, 4, or 5 amino acids, in either direction. Additional permissive sites can be identified using methods such as transposon mutagenesis (see, e.g., WO 02/102828 and WO 2008/036146). The insertions can be made at the indicated amino acids by insertion just C-terminal to the indicated amino acids (i.e., between amino acids 51-52, 207-208, 231-232, 277-278, 287-288, 340-341, and 436-437), or in place of short deletions (e.g., deletions of 1, 2, 3, 4, 5, 6, 7, or 8 amino acids) beginning at the indicated amino acids (or within 1-5 positions thereof, in either direction).

In addition to the envelope protein, insertions can be made into other virus proteins including, for example, the membrane/pre-membrane protein and NS1 (see, e.g., WO 2008/036146). For example, insertions can be made into a sequence preceding the capsid/pre-membrane cleavage site (at, e.g., -4, -2, or -1) or within the first 50 amino acids of the pre-membrane protein (e.g., at position 26), and/or between amino acids 236 and 237 of NS1 (or in regions surrounding the indicated sequences, as described above). In other examples, insertions can be made intergenically. For example, an insertion can be made between E and NS1 proteins and/or between NS2B and NS3 proteins (see, e.g., WO 2008/100464). In one example of an intergenic insertion, the inserted sequence can be fused with the C-terminus of the E protein of the vector, after the C-terminal signal/anchor sequence of the E protein, and the insertion can include a C-terminal anchor/signal sequence, which is fused with vector NS1 sequences. In another example of an intergenic insertion, the inserted sequences, with flanking protease cleavage

sites (e.g., YF 17D cleavage sites), can be inserted into a unique restriction site introduced at the NS2B/NS3 junction (WO 2008/100464).

In other examples, a sequence can be inserted in the context of an internal ribosome entry site (IRES, e.g., an IRES derived from encephalomyocarditis virus; EMCV), as noted above, such as inserted in the 3'-untranslated region (WO 2008/094674). In one example of such a vector, employing, for example, yellow fever virus sequences, an IRES-immunogen cassette can be inserted into a multiple cloning site engineered into the 3'-untranslated region of the vector, e.g., in a deletion (e.g., a 136 nucleotide deletion in the case of a yellow fever virus-based example) after the polyprotein stop codon (WO 2008/094674).

Details concerning the insertion of rabies virus G protein and full-length respiratory syncytial virus (RSV) F protein into s-PIV and d-PIV vectors of the invention are provided below in Example 3. The information provided in Example 3 can be applied in the context of other vectors and immunogens described herein.

Immunogens

PIVs (s-PIVs and d-PIVs) based on flavivirus sequences and live, attenuated chimeric flaviviruses (e.g., YF/TBE, YF/LGT, WN/TBE, and WN/LGT), as described above, can be used in the invention to deliver foreign (e.g., non-flavivirus) pathogen (e.g., viral, bacterial, fungal, and parasitic pathogens), cancer, and allergy-related immunogens. As discussed further below, in certain examples, it may be advantageous to target several pathogens occupying the same ecological niche, in a particular geographical region. Specific, non-limiting examples of such immunogens are provided as follows.

In addition to TBE virus, ticks are known to transmit another major disease, Lyme disease. Thus, in a first example, PIVs of the invention, such as PIVs including TBE/LGT sequences, as well as chimeric flaviviruses including TBE sequences (e.g., YF/TBE, YF/LGT, WN/TBE, LGT/TBE, and WN/LGT; in all instances where "TBE" is indicated, this includes the option of using the Hypr strain), can be used as vectors to deliver protective immunogens of the causative agent of Lyme disease (tick-borne spirochete *Borrelia burgdorferi*). This combination, targeting both infectious agents (TBE and *B. burgdorferi*) is advantageous, because TBE and Lyme disease are both tick-borne diseases. The PIV approaches can be applied to chimeras (e.g., YF/TBE, YF/LGT, WN/TBE, or WN/LGT), according to the invention, as well as to non-chimeric TBE and LGT viruses. An exemplary immunogen from *B. burgdorferi* that can be used in the invention is OspA (Gipson et al., Vaccine 21:3875-3884, 2003). Optionally, to increase safety and/or immunogenicity, OspA can be mutated to reduce chances of autoimmune responses and/or to eliminate sites for unwanted post-translational modification in vertebrate animal cells, such as N-linked glycosylation, which may affect immunogenicity of the expression product. Mutations that decrease autoimmunity can include, e.g., those described by Willett et al., Proc. Natl. Acad. Sci. U.S.A. 101:1303-1308, 2004. In one example, FTK-OspA, a putative cross-reactive T cell epitope, Bb OSpA₁₆₅₋₁₇₃ (YVLEGTLLTA) is altered to resemble the corresponding peptide sequence of *Borrelia afzelli* (FTLEGGKLVAN). In FTK-OspA, the corresponding sequence is FTLEGGKLLTA.

The sequence of OspA is as follows:

```

1 mkkyllylgil ilaliackgn vssldeknsv svdlpgemkv lvskeknkdg
kydliatvdk
61 lelkgtsdkn ngsgvlegvk adkskvklti sddlqgttle vfkedgkltv
skkvtskdks
121 steekfnekv evsekiitra dgtrleyttgi ksdgsgkake vlkgyvlegt
ltaekttlvv
181 kegtvtlslkn isksgevsve lndtdssaak khtaawnsqt stltitvnsk
ktdkdlvftke
241 ntitvqgyds ngtklegsav eitkldeikn alk

```

The full-length sequence and/or immunogenic fragments of the full-length sequence can be used in the present invention. Exemplary fragments can include one or more of domains 1 (amino acids 34-41), 2 (amino acids 65-75), 3 (amino acids 190-220), and 4 (amino acids 250-270) (Jiang et al., Clin. Diag. Lab. Immun. 1(4):406-412, 1994). Thus, for example, a peptide comprising any one (or more) of the following

```

64TRP (AF469170)
MKAFFVLSLL STAALTNAAR AGRLGSDLDL FGRVHGNYLA GIERAGPRGY PGLTASIGGE
VGARLGGRRAG VGVSSYGYGY PSWGYPYGGY GGYGGYGGYG GYDQGFSGAY GGYPGYGYGY
YPSGYGGGYG GSYGGSYGGG YTYPNVRASA GAAA

Isac (AF270496)
MRTAFTCALL AISFLGSPCS SSEDGLEQDT IVETTTQNLV ERHYRNSHGL CGAQYRNSSH
AEAVYNCTLN HLPVNVNATW EGIRHRINKT IPQFVKLICN FTVAMPQEFY LVYMGSDGNS
DFEEDKESTG TDEDSNTGSS AAAKVTEALI IEAENCTAH ITGWTTEPT TLEPTTESQF EAIP

Sal20 (EU008559)
MRTALTCALL AISFLGSPCS SSEGGLEKDS RVETTTQNLV ERYRKHPLGL CGAQYRNSSH
AEAVYNCTLS LLPLSVNTTW EGIRHRINKT IPEFVNLICN FTVAMPQFVY LVYMGSDGNS
YSEEDDGKT GSSAAVQVTE QLIIQAEENC TAHITGWTE APTTLEPTTE TQFEAIS

```

sequences (which include sequence variations that can be included in the sequence listed above, in any combination) can be delivered: LPGE/GM/IK/T/GVL; GTSDKN/S/DNGSGV/T; N/H/EIS/P/L/A/SK/NSGEV/IS/TV/AE/ALN/DDT/SD/NS/TS/TA/Q/RATKKA/GA/K/TWN/DS/AG/N/KT; SN/AGTK/NLEGS/N/K/TAVEIT/KK/TLD/KEI/LKN.

In addition to *B. burgdorferi* immunogens, tick saliva proteins, such as 64TRP, Isac, and Salp20, can be expressed, e.g., to generate a vaccine candidate of trivalent-specificity (TBE+ Lyme disease+ticks). Alternatively, tick saliva proteins can be expressed instead of *B. burgdorferi* immunogens in TBE sequence-containing vectors. In addition, there are many other candidate tick saliva proteins that can be used for tick vector vaccine development according to the invention (Francischetti et al., Insect Biochem. Mol. Biol. 35:1142-1161, 2005). One or more of these immunogens can be expressed in s-PIV-TBE. However, d-PIV-TBE may also be selected, because of its large insertion capacity. In addition to PIV-TBE, other PIV vaccines can be used as vectors, e.g., to protect from Lyme disease and another flavivirus disease, such as West Nile virus. Expression of these immunogens can be evaluated in cell culture, and immunogenicity/protection examined in available animal models (e.g., as described in

Gipson et al., Vaccine 21:3875-3884, 2003; Labuda et al., Pathog. 2(e27):0251-0259, 2006). Immunogens of other pathogens can be similarly expressed, in addition to Lyme disease and tick immunogens, with the purpose of making multivalent vaccine candidates. Exemplary tick saliva immunogens that can be used in the invention include the following:

Additional details concerning the TBE-related PIVs and LAVs are provided in Example 2, below.

The invention further provides PIV and LAV-vectored vaccines against other non-flavivirus pathogens, including vaccines having dual action, eliciting protective immunity against both flavivirus (as specified by the vector envelope proteins) and non-flavivirus pathogens (as specified by expressed immunologic determinant(s)). These are similar to the example of PIV-TBE-Lyme disease-tick vector vaccines described above. As mentioned above, such dual-action vaccines can be developed against a broad range of pathogens by expression of immunogens from, for example, viral, bacterial, fungal, and parasitic pathogens, and immunogens associated with cancer and allergy. As specific non-limiting examples, we describe herein the design and biological properties of PIV vectored-rabies and -respiratory syncytial virus (RSV) vaccine candidates constructed by expression of rabies virus G protein or full-length RSV F protein in place of or in combination with various deletions in one- and two-component PIV vectors (see Example 3, below). Also described in Example 4 are SIV/HIV-based PIV vectors. Example 5 provides influenza virus HA-based PIV vectors.

As is demonstrated in the Examples, below, s-PIV constructs may be advantageously used to stably deliver rela-

tively short foreign immunogens (similar to Lyme disease agent OspA protein and tick saliva proteins), because insertions are combined with a relatively short ΔC deletion. Two-component PIV vectors may be advantageously used to stably express relatively large immunogens, such as rabies G protein and RSV F, as the insertions in such vectors are combined with, for example, large $\Delta prM-E$, $\Delta C-prM-E$, and/or $\Delta NS1$ deletions. Some of the d-PIV components can be manufactured and used as vaccines individually, for instance, the PIV-RSV F construct described below containing a $\Delta C-prM-E$ deletion. In this case, the vaccine induces an immune response (e.g., neutralizing antibodies) predominantly against the expressed protein, but not against the flavivirus vector virus pathogen. In other examples of the invention, dual immunity is obtained by having immunity induced both to vector and insert components. Additionally, because of the large insertion capacity of PIV vectors, and the option of using two-component genomes, PIV vectors offer the opportunity to target several non-flavivirus pathogens simultaneously, e.g., by expressing foreign immunogens from two different non-flavivirus pathogens in the two components of a d-PIV.

In addition to the RSV F protein, rabies G protein, Lyme disease protective immunogens, and tick saliva proteins, as examples of foreign immunogens described above, other foreign immunogens can be expressed to target respective diseases including, for example, influenza virus type A and B immunogens. In these examples, a few short epitopes and/or whole genes of viral particle proteins can be used, such as the M2, HA, and NA genes of influenza A, and/or the NB or BM2 genes of influenza B (see, e.g., the PIV constructs of Example 5 below). Shorter fragments of M2, NB, and BM2, corresponding for instance to M2e, the extracellular fragment of M2, can also be used. In addition, fragments of the HA gene, including epitopes identified as HA0 (23 amino acids in length, corresponding to the cleavage site in HA) can be used. Specific examples of influenza-related sequences that can be used in the invention include PAKLLKERGFFGAIAGFLE (HA0), PAKLLKERGFFGAIAGFLESGC (HA0), NNAT-FNYTNPISHIRGS (NBe), MSLLETEVET-PIRNEWGCRCNDSSD (M2e), MSLLETEVETPTRNEWECRCSDDSSD (M2e), MSLLETEVETLTRNGWGCRCSDSSD (M2e), EVETPTRN (M2e), SLLTEVET-PIRNEWGCRCNDSSD (M2e), and SLLTEVET-PIRNEWGCR (M2e). Additional M2e sequences that can be used in the invention include sequences from the extracellular domain of BM2 protein of influenza B (consensus MLEPFQ, e.g., LEPFQILSISGC), and the M2e peptide from the H5N1 avian flu (MSLLETEVETLTRNGWGCRCSDSSD).

Other examples of pathogen immunogens that can be delivered in the vectors of the invention include codon-optimized SIV or HIV gag (55 kDa), gp120, gp140, gp145, gp41, gp160, SIV mac239 pol/-rev/tat/nef/pro genes or analogs or homologs and/or other naturally occurring variants from SIV and/or HIV, and other SIV and/or HIV immunogens (see, e.g., the PIV vectors described in Example 4 below); immunogens from HPV viruses, such as HPV16, HPV18, etc., e.g., the capsid protein L1 which self-assembles into HPV-like particles, the capsid protein L2 or its immunodominant portions (e.g., amino acids 1-200, 1-88, or 17-36), the E6 and E7 proteins which are involved in transforming and immortalizing mammalian cells fused together and appropriately mutated (fusion of the two genes creates a fusion protein, referred to as E6E7Rb⁻, that is about 10-fold less capable of transforming fibroblasts, and mutations of the E7 component at 2 residues renders the resulting fusion protein mutant incapable of inducing transformation (Boursnell et al., Vaccine

14:1485-1494, 1996). Other immunogens include protective immunogens from HCV, CMV, HSV2, viruses, malaria parasite, *Mycobacterium tuberculosis* causing tuberculosis, *C. difficile*, and other nosocomial infections, that are known in the art, as well as fungal pathogens, cancer immunogens, and proteins associated with allergy that can be used as vaccine targets.

Foreign immunogen inserts of the invention can be modified in various ways. For instance, codon optimization is used to increase the level of expression and eliminate long repeats in nucleotide sequences to increase insert stability in the RNA genome of PIV vectors. Immunogenicity can be increased by chimerization of proteins with immunostimulatory moieties well known in the art, such as TLR agonists, stimulatory cytokines, components of complement, heat-shock proteins, etc. (e.g., reviewed in "Immunopotentiators in Modern Vaccines," Schijns and O'Hagan Eds., 2006, Elsevier Academic Press: Amsterdam, Boston).

With respect to construction of dual vaccines against rabies and other flavivirus diseases, other combinations, such as TBE+rabies, YF+rabies, etc., can be of interest both for human and veterinary use in corresponding geographical regions, and thus can be similarly generated. Possible designs of expression constructs are not limited to those described herein. For example deletions and insertions can be modified, genetic elements can be rearranged, or other genetic elements (e.g. non-flavivirus, non-rabies signals for secretion, intracellular transport determinants, inclusion of or fusion with immunostimulatory moieties such as cytokines, TLR agonists such as flagellin, multimerization components such as leucine zipper, and peptides that increase the period of protein circulation in the blood) can be used to facilitate antigen presentation and increase immunogenicity. Further, such designs can be applied to s-PIV and d-PIV vaccine candidates based on vector genomes of other flaviviruses, and expressing immunogens of other pathogens, e.g., including but not limited to pathogens described in elsewhere herein.

Other examples of PIV and LAV vectors of the invention including combination vaccines such as DEN+Chikungunya virus (CHIKV) and YF+CHIKV. CHIKV, an alphavirus, is endemic in Africa, South East Asia, Indian subcontinent and the Islands, and the Pacific Islands and shares ecological/geographical niches with YF and DEN1-4. It causes serious disease primarily associated with severe pain (arthritis, other symptoms similar to DEN) and long-lasting sequelae in the majority of patients (Simon et al., Med. Clin. North Am. 92:1323-1343, 2008; Seneviratne et al., J. Travel Med. 14:320-325, 2007). Other examples of PIV and LAV vectors of the invention include YF+Ebola or DEN+Ebola, which co-circulate in Africa.

Immunogens for the above-noted non-flavivirus pathogens, sequences of which are well known in the art, may include glycoprotein B or a pp65/IE1 fusion protein of CMV (Reap et al., Vaccine 25(42):7441-7449, 2007; and references therein), several TB proteins (reviewed in Skeiky et al., Nat. Rev. Microbiol. 4(6):469-476, 2006), malaria parasite antigens such as RTS,S (a pre-erythrocytic circumsporozoite protein, CSP) and others (e.g., reviewed in Li et al., Vaccine 25(14):2567-2574, 2007), CHIKV envelope proteins E1 and E2 (or the C-E2-E1, E2-E1 cassettes), HCV structural proteins C-E1-E2 forming VLPs (Ezelle et al., J. Virol. 76(23):12325-12334, 2002) or other proteins to induce T-cell responses, Ebola virus glycoprotein GP (Yang et al., Virology 377(2):255-264, 2008).

In addition to the immunogens described above, the vectors described herein may include one or more immunogen(s) derived from or that direct an immune response against one or

more viruses (e.g., viral target antigen(s)) including, for example, a dsDNA virus (e.g., adenovirus, herpesvirus, epstein-barr virus, herpes simplex type 1, herpes simplex type 2, human herpes virus simplex type 8, human cytomegalovirus, varicella-zoster virus, poxvirus); ssDNA virus (e.g., parvovirus, papillomavirus (e.g., E1, E2, E3, E4, E5, E6, E7, E8, BPV1, BPV2, BPV3, BPV4, BPV5, and BPV6 (In Papillomavirus and Human Cancer, edited by H. Pfister (CRC Press, Inc. 1990)); Lancaster et al., Cancer Metast. Rev. pp. 6653-6664, 1987; Pfister et al., Adv. Cancer Res. 48:113-147, 1987)); dsRNA viruses (e.g., reovirus); (+)ssRNA viruses (e.g., picornavirus, coxsackie virus, hepatitis A virus, poliovirus, togavirus, rubella virus, flavivirus, hepatitis C virus, yellow fever virus, dengue virus, west Nile virus); (-)ssRNA viruses (e.g., orthomyxovirus, influenza virus, rhabdovirus, paramyxovirus, measles virus, mumps virus, parainfluenza virus, rhabdovirus, rabies virus); ssRNA-RT viruses (e.g., retrovirus, human immunodeficiency virus (HIV)); and dsDNA-RT viruses (e.g. hepadnavirus, hepatitis B). Immunogens may also be derived from other viruses not listed above but available to those of skill in the art.

With respect to HIV, immunogens may be selected from any HIV isolate. As is well-known in the art, HIV isolates are now classified into discrete genetic subtypes. HIV-1 is known to comprise at least ten subtypes (A, B, C, D, E, F, G, H, J, and K). HIV-2 is known to include at least five subtypes (A, B, C, D, and E). Subtype B has been associated with the HIV epidemic in homosexual men and intravenous drug users worldwide. Most HIV-1 immunogens, laboratory adapted isolates, reagents and mapped epitopes belong to subtype B. In sub-Saharan Africa, India, and China, areas where the incidence of new HIV infections is high, HIV-1 subtype B accounts for only a small minority of infections, and subtype HIV-1 C appears to be the most common infecting subtype. Thus, in certain embodiments, it may be desirable to select immunogens from HIV-1 subtypes B and/or C. It may be desirable to include immunogens from multiple HIV subtypes (e.g., HIV-1 subtypes B and C, HIV-2 subtypes A and B, or a combination of HIV-1 and HIV-2 subtypes) in a single immunological composition. Suitable HIV immunogens include ENV, GAG, PRO, POL, NEF, as well as variants, derivatives, and fusion proteins thereof, for example.

Further, as described in Example 4 in reference to particular constructs, the invention includes constructs including multiple different proteins in a single precursor, wherein the open reading frames may be, optionally, separated by protease cleavage sites, such as FMDV 2A cleavage sites, as described herein. Thus, in one example, a cassette may include gp120 (e.g., modified as described in Example 4), gag, and pro genes from SIV or HIV. Further, the invention includes the hybrid sequences including, e.g., heterologous transmembrane and/or signal sequences, as described in detail in Example 4. Thus, for example, the invention includes the use of rabies virus G protein-specific signals and/or anchor sequences in the context of gp120-containing PIV constructs, as described herein.

Immunogens may also be derived from or direct an immune response against one or more bacterial species (spp.) (e.g., bacterial target antigen(s)) including, for example, *Bacillus* spp. (e.g., *Bacillus anthracis*), *Bordetella* spp. (e.g., *Bordetella pertussis*), *Borrelia* spp. (e.g., *Borrelia burgdorferi*), *Brucella* spp. (e.g., *Brucella abortus*, *Brucella canis*, *Brucella melitensis*, *Brucella suis*), *Campylobacter* spp. (e.g., *Campylobacter jejuni*), *Chlamydia* spp. (e.g., *Chlamydia pneumoniae*, *Chlamydia psittaci*, *Chlamydia trachomatis*), *Clostridium* spp. (e.g., *Clostridium botulinum*, *Clostridium difficile*, *Clostridium perfringens*, *Clostridium*

tetani), *Corynebacterium* spp. (e.g., *Corynebacterium diphtheriae*), *Enterococcus* spp. (e.g., *Enterococcus faecalis*, *enterococcus faecum*), *Escherichia* spp. (e.g., *Escherichia coli*), *Francisella* spp. (e.g., *Francisella tularensis*), *Haemophilus* spp. (e.g., *Haemophilus influenza*), *Helicobacter* spp. (e.g., *Helicobacter pylori*), *Legionella* spp. (e.g., *Legionella pneumophila*), *Leptospira* spp. (e.g., *Leptospira interrogans*), *Listeria* spp. (e.g., *Listeria monocytogenes*), *Mycobacterium* spp. (e.g., *Mycobacterium leprae*, *Mycobacterium tuberculosis*), *Mycoplasma* spp. (e.g., *Mycoplasma pneumoniae*), *Neisseria* spp. (e.g., *Neisseria gonorrhoea*, *Neisseria meningitidis*), *Pseudomonas* spp. (e.g., *Pseudomonas aeruginosa*), *Rickettsia* spp. (e.g., *Rickettsia rickettsii*), *Salmonella* spp. (e.g., *Salmonella typhi*, *Salmonella typhimurium*), *Shigella* spp. (e.g., *Shigella sonnei*), *Staphylococcus* spp. (e.g., *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Staphylococcus saprophyticus*, coagulase negative staphylococcus (e.g., U.S. Pat. No. 7,473,762)), *Streptococcus* spp. (e.g., *Streptococcus agalactiae*, *Streptococcus pneumoniae*, *Streptococcus pyrogenes*), *Treponema* spp. (e.g., *Treponema pallidum*), *Vibrio* spp. (e.g., *Vibrio cholerae*), and *Yersinia* spp. (*Yersinia pestis*). Immunogens may also be derived from or direct the immune response against other bacterial species not listed above but available to those of skill in the art.

Immunogens may also be derived from or direct an immune response against one or more parasitic organisms (spp.) (e.g., parasite target antigen(s)) including, for example, *Ancylostoma* spp. (e.g., *A. duodenale*), *Anisakis* spp., *Ascaris lumbricoides*, *Balantidium coli*, *Cestoda* spp., *Cimicidae* spp., *Clonorchis sinensis*, *Dicrocoelium dendriticum*, *Dicrocoelium hospes*, *Diphyllobothrium latum*, *Dracunculus* spp., *Echinococcus* spp. (e.g., *E. granulosus*, *E. multilocularis*), *Entamoeba histolytica*, *Enterobius vermicularis*, *Fasciola* spp. (e.g., *F. hepatica*, *F. magna*, *F. gigantica*, *F. jacksoni*), *Fasciolopsis buski*, *Giardia* spp. (*Giardia lamblia*), *Gnathostoma* spp., *Hymenolepis* spp. (e.g., *H. nana*, *H. diminuta*), *Leishmania* spp., *Loa loa*, *Metorchis* spp. (*M. conjunctus*, *M. albidus*), *Necator americanus*, *Oestroidea* spp. (e.g., botfly), *Onchocercidae* spp., *Opisthorchis* spp. (e.g., *O. viverrini*, *O. felineus*, *O. guayaquilensis*, and *O. noverca*), *Plasmodium* spp. (e.g., *P. falciparum*), *Protofasciola robusta*, *Parafasciolopsis fasciomorphae*, *Paragonimus westermani*, *Schistosoma* spp. (e.g., *S. mansoni*, *S. japonicum*, *S. mekongi*, *S. haematobium*), *Spirometra erinaceieuropaei*, *Strongyloides stercoralis*, *Taenia* spp. (e.g., *T. saginata*, *T. solium*), *Toxocara* spp. (e.g., *T. canis*, *T. cati*), *Toxoplasma* spp. (e.g., *T. gondii*), *Trichobilharzia regenti*, *Trichinella spiralis*, *Trichuris trichiura*, *Trombiculidae* spp., *Trypanosoma* spp., *Tunga penetrans*, and/or *Wuchereria bancrofti*. Immunogens may also be derived from or direct the immune response against other parasitic organisms not listed above but available to those of skill in the art.

Immunogens may be derived from or direct the immune response against tumor target antigens (e.g., tumor target antigens). The term tumor target antigen (TA) may include both tumor-associated antigens (TAAs) and tumor-specific antigens (TSAs), where a cancerous cell is the source of the antigen. A TA may be an antigen that is expressed on the surface of a tumor cell in higher amounts than is observed on normal cells or an antigen that is expressed on normal cells during fetal development. A TSA is typically an antigen that is unique to tumor cells and is not expressed on normal cells. TAs are typically classified into five categories according to their expression pattern, function, or genetic origin: cancer-testis (CT) antigens (i.e., MAGE, NY-ESO-1); melanocyte differentiation antigens (e.g., Melan A/MART-1, tyrosinase, gp100); mutational antigens (e.g., MUM-1, p53, CDK-4);

overexpressed 'self' antigens (e.g., HER-2/neu, p53); and viral antigens (e.g., HPV, EBV). Suitable TAs include, for example, gp100 (Cox et al., *Science* 264:716-719, 1994), MART-1/Melan A (Kawakami et al., *J. Exp. Med.*, 180:347-352, 1994), gp75 (TRP-1) (Wang et al., *J. Exp. Med.*, 186: 1131-1140, 1996), tyrosinase (Wolfel et al., *Eur. J. Immunol.*, 24:759-764, 1994), NY-ESO-1 (WO 98/14464; WO 99/18206), melanoma proteoglycan (Hellstrom et al., *J. Immunol.*, 130:1467-1472, 1983), MAGE family antigens (e.g., MAGE-1, 2, 3, 4, 6, and 12; Van der Bruggen et al., *Science* 254:1643-1647, 1991; U.S. Pat. No. 6,235,525), BAGE family antigens (Boel et al., *Immunity* 2:167-175, 1995), GAGE family antigens (e.g., GAGE-1.2; Van den Eynde et al., *J. Exp. Med.* 182:689-698, 1995; U.S. Pat. No. 6,013,765), RAGE family antigens (e.g., RAGE-1; Gaugler et al., *Immunogenetics* 44:323-330, 1996; U.S. Pat. No. 5,939,526), N-acetylglucosaminyltransferase-V (Guilloux et al., *J. Exp. Med.* 183:1173-1183, 1996), p15 (Robbins et al., *J. Immunol.* 154:5944-5950, 1995), β -catenin (Robbins et al., *J. Exp. Med.*, 183:1185-1192, 1996), MUM-1 (Coulie et al., *Proc. Natl. Acad. Sci. U.S.A.* 92:7976-7980, 1995), cyclin dependent kinase-4 (CDK4) (Wolfel et al., *Science* 269: 1281-1284, 1995), p21-ras (Fossum et al., *Int. J. Cancer* 56:40-45, 1994), BCR-abl (Bocchia et al., *Blood* 85:2680-2684, 1995), p53 (Theobald et al., *Proc. Natl. Acad. Sci. U.S.A.* 92:11993-11997, 1995), p185 HER2/neu (erb-B1; Fisk et al., *J. Exp. Med.*, 181:2109-2117, 1995), epidermal growth factor receptor (EGFR) (Harris et al., *Breast Cancer Res. Treat.* 29:1-2, 1994), carcinoembryonic antigens (CEA) (Kwong et al., *J. Natl. Cancer Inst.*, 85:982-990, 1995) U.S. Pat. Nos. 5,756,103; 5,274,087; 5,571,710; 6,071,716; 5,698,530; 6,045,802; EP 263933; EP 346710; and EP 784483; carcinoma-associated mutated mucins (e.g., MUC-1 gene products; Jerome et al., *J. Immunol.*, 151:1654-1662, 1993); EBNA gene products of EBV (e.g., EBNA-1; Rickinson et al., *Cancer Surveys* 13:53-80, 1992); E7, E6 proteins of human papillomavirus (Ressing et al., *J. Immunol.* 154:5934-5943, 1995); prostate specific antigen (PSA; Xue et al., *The Prostate* 30:73-78, 1997); prostate specific membrane antigen (PSMA; Israeli et al., *Cancer Res.* 54:1807-1811, 1994); idiotypic epitopes or antigens, for example, immunoglobulin idiotypes or T cell receptor idiotypes (Chen et al., *J. Immunol.* 153:4775-4787, 1994); KSA (U.S. Pat. No. 5,348,887), kinesin 2 (Dietz, et al., *Biochem. Biophys. Res. Commun.* 275(3): 731-738, 2000), HIP-55, TGF β -1 anti-apoptotic factor (Toomey et al., *Br. J. Biomed. Sci.* 58(3):177-183, 2001), tumor protein D52 (Bryne et al., *Genomics* 35:523-532, 1996), H1FT, NY-BR-1 (WO 01/47959), NY-BR-62, NY-BR-75, NY-BR-85, NY-BR-87, and NY-BR-96 (Scanlan, M. Serologic and Bioinformatic Approaches to the Identification of Human Tumor Antigens, in *Cancer Vaccines 2000*, Cancer Research Institute, New York, N.Y.), and/or pancreatic cancer antigens (e.g., SEQ ID NOs: 1-288 of U.S. Pat. No. 7,473,531). Immunogens may also be derived from or direct the immune response against include TAs not listed above but available to one of skill in the art.

In addition to the specific immunogen sequences listed above, the invention also includes the use of analogs of the sequences. Such analogs include sequences that are, for example, at least 80%, 90%, 95%, or 99% identical to the reference sequences, or fragments thereof. The analogs also include fragments of the reference sequences that include, for example, one or more immunogenic epitopes of the sequences. Further, the analogs include truncations or expansions of the sequences (e.g., insertion of additional/repeat immunodominant/helper epitopes) by, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11-20, etc., amino acids on either or both ends.

Truncation may remove immunologically unimportant or interfering sequences, e.g., within known structural/immunologic domains, or between domains; or whole undesired domains can be deleted; such modifications can be in the ranges 21-30, 31-50, 51-100, 101-400, etc. amino acids. The ranges also include, e.g., 20-400, 30-100, and 50-100 amino acids.

Cocktails

The invention also includes compositions including mixtures of two or more PIVs and/or PIV vectors, as described herein. As discussed above, use of such mixtures or cocktails may be particularly advantageous when induction of immunity to more than one immunogen and/or pathogen is desired. This may be useful, for example, in vaccination against different flaviviruses that may be endemic to the region in which the vaccine recipient resides. This may also be useful in the context of administration of multiple immunogens against the same target.

Non-limiting examples of PIV cocktails included in the invention are those including PIV-JE+PIV-DEN, and PIV-YF+PIV-DEN. In both of these examples, the PIVs for either or both components can be single or dual component PIVs, as described above. In addition, in the case of the PIV-DEN, the PIV can include sequences of just one dengue serotype selected from the group consisting of dengue serotypes 1-4, or the cocktail can include PIVs expressing sequences from two, three, or all four of the serotypes. Further, the TBE/*Borrelia burgdorferi*/tick saliva protein (e.g., 64TRP, Isac, Salp20) vaccines described herein can be based on including the different immunogens within a single PIV or live attenuated flavivirus, or can be based on mixtures of PIVs (or LAVs), which each include one or more of the immunogens. The cocktails of the invention can be formulated as such or can be mixed just prior to administration.

Use, Formulation, and Administration

The invention includes the PIV vectors, PIVs, LAV vectors, and LAVs, as well as corresponding nucleic acid molecules, pharmaceutical or vaccine compositions, and methods of their use and preparation. The PIV vectors, PIVs, LAV vectors, and LAVs of the invention can be used, for example, in vaccination methods to induce an immune response to TBE or other flavivirus, and/or another expressed immunogen, as described herein. These methods can be prophylactic, in which case they are carried out on subjects (e.g., human subjects or other mammalian subjects) not having, but at risk of developing infection or disease caused by TBE or another flavivirus and/or a pathogen from which the other expressed immunogen is derived. The methods can also be therapeutic, in which they are carried out on subjects already having an infection by one or more of the relevant pathogens. Further, the viruses and vectors can be used individually or in combination with one another or other vaccines. The subjects treated according to the methods of the invention include humans, as well as non-human mammals (e.g., livestock, such as, cattle, pigs, horses, sheep, and goats, and domestic animals, including dogs and cats).

Formulation of the PIV vectors, PIVs, LAV vectors, and LAVs of the invention can be carried out using methods that are standard in the art. Numerous pharmaceutically acceptable solutions for use in vaccine preparation are well known and can readily be adapted for use in the present invention by those of skill in this art (see, e.g., *Remington's Pharmaceutical Sciences* (18th edition), ed. A. Gennaro, 1990, Mack Publishing Co., Easton, Pa.). In two specific examples, the PIV vectors, PIVs, LAV vectors, and LAVs are formulated in Minimum Essential Medium Earle's Salt (MEME) containing 7.5% lactose and 2.5% human serum albumin or MEME

containing 10% sorbitol. However, the PIV vectors, PIVs, LAV vectors, and LAVs can simply be diluted in a physiologically acceptable solution, such as sterile saline or sterile buffered saline.

The PIV vectors, PIVs, LAV vectors, and LAVs of the invention can be administered using methods that are well known in the art, and appropriate amounts of the viruses and vectors to be administered can readily be determined by those of skill in the art. What is determined to be an appropriate amount of virus to administer can be determined by consideration of factors such as, e.g., the size and general health of the subject to whom the virus is to be administered. For example, in the case of live, attenuated viruses of the invention, the viruses can be formulated as sterile aqueous solutions containing between 10^2 and 10^8 , e.g., 10^3 to 10^7 , infectious units (e.g., plaque-forming units or tissue culture infectious doses) in a dose volume of 0.1 to 1.0 ml. PIVs can be administered at similar doses and in similar volumes; PIV titers however are usually measured in, e.g., focus-forming units determined by immunostaining of foci, as these defective constructs tend not to form virus-like plaques. Doses can range between 10^2 and 10^8 FFU and administered in volumes of 0.1 to 1.0 ml.

All viruses and vectors of the invention can be administered by, for example, intradermal, subcutaneous, intramuscular, intraperitoneal, or oral routes. In specific examples, dendritic cells are targeted by intradermal or transcutaneous administration, by use of, for example, microneedles or microabrasion devices. Further, the vaccines of the invention can be administered in a single dose or, optionally, administration can involve the use of a priming dose followed by a booster dose that is administered, e.g., 2-6 months later, as determined to be appropriate by those of skill in the art. Optionally, PIV vaccines can be administered via DNA or RNA immunization using methods known to those skilled in the art (Chang et al., Nat. Biotechnol. 26:571-577, 2008; Kofler et al., Proc. Natl. Acad. Sci. U.S.A. 101:1951-1956, 2004).

Optionally, adjuvants that are known to those skilled in the art can be used in the administration of the viruses and vectors of the invention. Adjuvants that can be used to enhance the immunogenicity of the viruses include, for example, liposomal formulations, synthetic adjuvants, such as (e.g., QS21), muramyl dipeptide, monophosphoryl lipid A, polyphosphazine, CpG oligonucleotides, or other molecules that appear to work by activating Toll-like Receptor (TLR) molecules on the surface of cells or on nuclear membranes within cells. Although these adjuvants are typically used to enhance immune responses to inactivated vaccines, they can also be used with live or replication-defective vaccines. Both agonists of TLRs or antagonists may be useful in the case of live or replication-defective vaccines. The vaccine candidates can be designed to express TLR agonists. In the case of a virus delivered via a mucosal route, for example, orally, mucosal adjuvants such as the heat-labile toxin of *E. coli* (LT) or mutant derivations of LT can be used as adjuvants. In addition, genes encoding cytokines that have adjuvant activities can be inserted into the vaccine candidates. Thus, genes encoding desired cytokines, such as GM-CSF, IL-2, IL-12, IL-13, IL-5, etc., can be inserted together with foreign immunogen genes to produce a vaccine that results in enhanced immune responses, or to modulate immunity directed more specifically towards cellular, humoral, or mucosal responses (e.g., reviewed in "Immunopotentiators in Modern Vaccines", Schijns and O'Hagan Eds., 2006, Elsevier Academic Press: Amsterdam, Boston, etc.). Optionally, a patch containing a layer of an appropriate toxin-derived adjuvant, can be

applied over the injection site. Toxin promotes local inflammation attracting lymphocytes, which leads to a more robust immune response.

EXAMPLES

Additional details concerning the invention are provided in the Examples, below. In the Examples, experiments are described in which PIVs based on WN, JE, and YF viruses (see, e.g., WO 2007/098267 and WO 2008/137163) were tested. Firstly, we demonstrated that the constructs are significantly more attenuated in a sensitive suckling mouse neurovirulence model (zero mortality at all tested doses) as compared to available LAV controls (YF 17D, YF/JE LAV, and YF/WN LAV). We demonstrated for the first time that d-PIV constructs were avirulent in this model and thus that two-component PIVs do not undergo uncontrolled (unlimited) spread in vivo and cannot cause clinical signs. Secondly, we performed comparisons of the immunogenicity and efficacy of the PIVs and the LAVs, and demonstrated that PIV vaccines can induce immune response comparable to LAVs and be equally efficacious (e.g., as observed for PIV-WN and YF/WN LAV pair of vaccines). In one pair examined, YF 17D LAV was significantly more immunogenic than PIV-YF. Thus, production of VLPs can vary between different, similarly designed PIV constructs. Specifically, we propose that PIV-YF does not generate a large amount of YF VLPs compared to PIV-WN (WN VLPs), and that increased production of VLPs can be achieved by genetic modifications at the C/prM junction in suboptimal PIV constructs. Specifically, the C/prM junction is an important location in the flavivirus polyprotein orchestrating the formation of viral envelope and synthesis of viral proteins (Yamshchikov and Compans, Virology 192:38-51, 1993; Amberg and Rice, J. Virol. 73:8083-8094, 1999; Stocks and Lobigs, J. Virol. 72:2141-2149, 1998). We propose that secretion of VLPs in PIV infected cells (in contrast to production of viral particles in whole viruses) can be increased by uncoupling of the viral protease and signalase cleavages at the junction, or use of a strong heterologous signal peptide (tPA, etc.) in place of the signal for prM, or by mutagenesis of the signal for prM. The efficiency of signalase cleavage at the C/prM junction of flaviviruses is low (Stocks and Lobigs, J. Virol. 72:2141-2149, 1998), e.g., as predicted by SignalP 3.0 on-line program. It is expected that more efficient cleavage efficiency can be achieved by analysis of specific amino acid substitutions near the cleavage site with SignalP 3.0 (e.g., as described in application WO 2008/100464), followed by incorporation of chosen mutation(s) into PIV genomes, recovery of PIV progeny and measuring VLP secretion. Non-flavivirus signals are inserted by methods standard in the art. Uncoupling between the viral protease and signalase cleavages can be achieved by ablating the viral cleavage site by any non-conservative mutation (e.g., RRS in YF17D C to RRA or GRS or RSS, etc.), or deletion of the entire site or some of its 3 residues. If necessary, formation of free N-terminus of the signal of foreign protein can be achieved by using such elements as autoprotease, or termination codon followed by an IRES. Alternatively, the native AUG initiation codon of C can be ablated (in constructs where C protein sequence is unnecessary, e.g., Δ C PIV) and AUG placed in front of foreign gene. Optimization of vector signal can be performed by random mutagenesis, e.g., by insertion of synthetic randomized sequence followed by identification of viable PIV variants with increased VLP secretion.

We also discovered that PIV constructs were substantially more immunogenic in hamsters when administered by the IP

route, as compared to the subcutaneous route. We concluded that this was most likely due to better targeting of antigen presenting cells in lymphoid tissues, which are abundant in the abdomen, but not abundant in tissues underlying the skin. Based on these observations, we concluded that efficient targeting of PIVs to dendritic cells, abundant in the skin, can be achieved by cutaneous inoculation, e.g., via skin microabrasion or intradermal injection using microneedles (Dean et al., *Hum Vaccin.* 1:106-111, 2005).

Further, we have carried out experiments to show the feasibility of administering mixtures, or cocktails, of different PIVs, such as those described herein (e.g., JE+DEN and YF+DEN). In order to administer cocktails, it is important to verify that there is no interference between co-administered components, and that a balanced immune response is induced. Several PIV mixtures were used to immunize rodents and immune responses were compared to PIV constructs administered individually. No interference was observed in mixtures, and thus cocktail PIV vaccines are feasible. Such formulations may be of particular significance in geographical regions where different flaviviruses co-circulate. This could be also used to simultaneously administer several PIV-based vaccines against non-flavivirus pathogens.

Further, we have demonstrated that no neutralizing antibody response is induced against packaging envelope after at least two doses of PIV (and thus antibodies are elicited against VLPs secreted from infected cells). This was demonstrated using the helper (AprM-E) component of a d-PIV (see in FIG. 2) packaged individually, or by measuring neutralizing antibodies to heterologous packaging envelopes (e.g., to the WN envelope used to package PIV-JE in helper cells providing WN-specific C-prM-E proteins in trans). The latter observations support sequential use of different PIV vaccines manufactured in a universal helper packaging cells line, and sequential use of different recombinant PIV-vectored vaccines in the same individual, as discussed above. In addition, we confirmed previous observations that PIV constructs can be stably propagated to high yields in vitro, and that no recombination restoring whole virus occurs after prolonged passaging in substrate cells (Mason et al., *Virology* 351:432-443, 2006; Shustov et al., *J. Virol.* 21:11737-11748, 2007).

These and other aspects of the invention are further described in the Examples, below.

Example 1

Pseudoinfectious Virus Platform Development Studies

Attenuation in Suckling Mouse Neurovirulence (NV) Model

Materials used in the studies described below are described in Table 1 and the references cited therein. These include s-PIV-WN (based on wt WN virus strain NY99 sequences), s-PIV-JE, s-PIV-WN/JE (based on wt WN virus backbone and prM-E genes from wt JE virus Nakayama strain), s-PIV-YF/WN (YF 17D backbone and prM-E genes from WN virus), and s-PIV-YF (based on YF 17D sequences). Additional materials include d-PIV-YF (YF d-PIV, grown in regular BHK cells (Shustov et al., *J. Virol.* 21:11737-11748, 2007), and two-component d-PIV-WN (grown in regular Vero cells; Suzuki et al., *J. Virol.* 82:6942-6951, 2008).

Attenuation of these PIV prototypes was compared to LAVs YF 17D, a chimeric YF/JE virus, and a chimeric YF/WN virus in suckling mouse NV test (IC inoculation) using highly susceptible 5-day old ICR mice (the chimeric viruses include yellow fever capsid and non-structural sequences, and JE or WN prM-E sequences). None of the

animals that received PIV constructs showed clinical signs or died, while mortality was observed in animals inoculated with LAVs (Table 2). The YF 17D virus is neurovirulent for mice of all ages, while the chimeric vaccines are not neurovirulent for adult mice, but can cause dose-dependent mortality in more sensitive suckling mice (Guirakhoo et al., *Virology* 257:363-372, 1999; Arroyo et al., *J. Virol.* 78:12497-12507, 2004). Accordingly, 90% of suckling mice that received doses as low as 1 PFU of YF 17D died. YF/JE and YF/WN LAVs caused partial mortality at much higher doses ($>2 \log_{10}$ PFU and $3 \log_{10}$ PFU, respectively), with longer average survival time (AST) of animals that died, as expected. Thus, PIV constructs are completely avirulent in this sensitive model (at least 20,000-200,000 times less neurovirulent than the licensed YF 17D vaccine).

The YF d-PIV and WN d-PIV caused no mortality or clinical signs. Thus, the two-component PIV variants that theoretically could spread within brain tissue from cells co-infected by both of their components did not cause disease. Moreover, we tried to detect the d-PIVs in the brains of additional animals in this experiment, sacrificed on day 6 post-inoculation by titration, and detected none (brain tissues from 10 and 11 mice that received $4 \log_{10}$ FFU of YF d-PIV and WN d-PIV, respectively, were homogenized and used for titration). Thus, the d-PIVs did not cause spreading infection characteristic of whole virus. YF/JE LAV has been shown to replicate in the brain of adult ICR mice inoculated by the IC route with a peak titer of $\sim 6 \log_{10}$ PFU/g on day 6, albeit without clinical signs (Guirakhoo et al., *Virology* 257:363-372, 1999). Co-infection of cells with components of a d-PIV is clearly a less efficient process than infection with whole virus. The data show that d-PIV replication in vivo is quickly brought under control by innate immune responses (and adaptive responses in older animals).

Immunogenicity/Efficacy in Mice and Hamsters

Immunogenicity/efficacy of the PIV prototypes described above was compared to that of chimeric LAV counterparts and YF 17D in mice and Syrian hamsters. The general experiment design is illustrated in FIG. 3 (mice, IP immunization). Experiments in hamsters were performed similarly (plus-minus a few days, SC or IP inoculation with doses indicated below). 3.5-week old ICR mice (for s-PIV-WN and -YF, YF/WN LAV, and YF 17D groups) or C57/BL6 mice (for s-PIV-JE and YF/JE LAV groups) were immunized IP with graded doses of PIV constructs ($4-6 \log_{10}$ FFU/dose) or chimeric LAV and YF 17D LAV controls ($4 \log_{10}$ PFU). Select PIV-WN, -JE and -YF groups were boosted on day 21 with $5 \log_{10}$ FFU of corresponding constructs (Table 3). Neutralizing antibody responses were determined in animal sera by standard PRNT₅₀ against YF/WN or /JE LAVs, or YF 17D viruses. PIV-WN induced very high WN-specific neutralizing antibody responses in all groups, with or without boost, as evidenced by PRNT₅₀ titers determined in pools of sera from immunized animals on days 20 and 34, which was comparable to that in the YF/WN LAV control group. Accordingly, animals immunized with both PIV-WN and YF/WN LAV were protected from lethal challenge on day 35 with wt WN virus (IP, 270 LD₅₀), but not mock-immunized animals (Table 3). When WN neutralizing antibodies were measured in sera from individual mice, high uniformity of immune responses was observed (FIG. 4). Thus, single-round PIV vaccines can be as immunogenic and efficacious as corresponding LAVs. PIV-JE was also highly immunogenic (black mice), while immunogenicity of PIV-YF was significantly lower compared to the YF 17D control (ICR mice). Yet, dose-dependent protection of PIV-YF immunized animals (but not mock-immunized animals) was observed following a severe lethal

IC challenge with wt YF strain Asibi virus (500 LD₅₀) (Table 3), which is in agreement with the knowledge that neutralizing antibody titers as low 1:10 are protective against flavivirus infections.

The YF 17D control virus was highly immunogenic (e.g., PRNT₅₀ titer 1:1,280 on day 34), and thus it is able to infect cells and replicate efficiently *in vivo*, and its envelope is a strong immunogen. Therefore, it is unlikely that low immunogenicity of PIV-YF was due to its inability to infect cells or replicate efficiently in infected cells *in vivo*. We believe that the low immunogenicity of PIV-YF (e.g., compared to PIV-WN) was most likely due to a low-level production of YF-specific VLPs in PIV-YF infected cells (while VLP secretion is high in PIV-WN infected cells). As discussed above, we propose that immunogenicity of PIV-YF can be significantly increased, e.g., by appropriate modifications at the C/prM junction, e.g., by uncoupling the two protease cleavages that occur at this junction (viral protease and signalase cleavages), and/or by using a strong heterologous signal [e.g., rabies virus G protein signal, or eukaryotic tissue plasminogen activator (tPA) signal (Malin et al., *Microbes and Infection*, 2:1677-1685, 2000), etc.] in place of the YF signal for prM.

A similar experiment was performed in ~4.5-week old Syrian hamsters, to compare immunogenicity of PIV constructs to LAV controls in this model. Animals were immunized SC with graded doses of the test articles (Table 4). PIV-WN was highly immunogenic, e.g., WN-specific PRNT₅₀ titers on day 38 (pre-challenge) were 1:320, 1:640, and 1:1280 in groups that received 5, 6, and 6 (prime)+5 (boost) log₁₀ FFU doses, respectively. This was somewhat lower compared to YF/WN LAV 4 log₁₀ PFU control (≥1:2560). PIV-JE and -YF induced detectable specific neutralizing antibody responses, albeit with lower titers compared to YF/JE LAV and YF 17D controls. All animals immunized with PIV-WN and YF/WN were solidly protected from lethal challenge with wt WN virus as evidenced by the absence of mortality and morbidity (e.g., loss of body weight after challenge), as well as absence or a significant reduction of post-challenge WN virus viremia. Mock-immunized animals were not protected (Table 4). PIV-JE and -WN protected animals from respective challenge in dose-dependent fashion. Protective efficacy in this experiment is additionally illustrated in FIG. 5. For example, high post-challenge YF virus (hamster adapted Asibi strain) viremia was observed in mock immunized animals, peaking on day 3 at a titer of >8 log₁₀ PFU/ml (upper left panel); all of the animals lost weight, and 1 out of 4 died (upper right panel). In contrast, viremia was significantly reduced or absent in hamsters immunized with PIV-YF (two doses; despite relatively low neutralizing titers) or YF 17D; none of these animals lost weight. Similarly, animals immunized with PIV-WN or YF/WN LAV were significantly or completely protected in terms of post-challenge WN virus viremia and body weight loss/mortality, in contrast to mock controls (compare in bottom panels). Thus, high immunogenicity/efficacy of PIV was demonstrated in a second animal model.

In another hamster experiment, animals were immunized with PIV constructs by the IP route, with two doses. Table 5 compares neutralizing immune responses (specific for each vaccine) determined in pooled sera of hamsters in the above-described experiment (SC inoculation) to those after IP immunization, for PIV-WN, -YF/WN, -WN/JE, and -YF after the first dose (days 20-21) and second dose (days 34-38). A clear effect of the immunization route was observed both after the 1st and 2nd doses. For instance, for PIV-WN after 1st dose, SC immunization resulted in WN-specific PRNT₅₀ titer of 1:40, while IP inoculation resulted in much higher titer 1:320

(and after the 2nd dose, titers were similar). A more pronounced effect was observed for other constructs after both the 1st and 2nd doses. Interestingly, PIV-YF/WN was very highly immunogenic by IP route (titer 1:320 after 1st IP dose vs. 1:20 by SC, and 1:1,280 after 2nd dose vs. 1:160 by SC). Similarly, immunogenicity of PIV-JE was significantly increased (e.g., JE-specific titer of 1:640 after two IP poses). Thus, better targeting of lymphoid cells, specifically antigen-presenting cells (which are more abundant in the abdomen as opposed to tissues under the skin), is an important consideration for use of PIV vaccines. In humans, efficient targeting of dendritic cells of the skin, increasing the magnitude of immune response, can be achieved by intradermal delivery, which we thus propose for a route for PIV immunization of humans.

In the above-described experiments, we also determined whether a neutralizing antibody response was induced against packaging envelopes (as opposed to response to VLPs encoded by PIV constructs and secreted by infected cells). No WN-specific neutralizing antibodies were detected by PRNT₅₀ in animals immunized with 5 log₁₀ FFU of the second component of WN d-PIV, containing the AC-prM-E deletion and thus not encoding VLPs, but packaged into the WN envelope in BHK-CprME(WN) helper cells, and no YF-specific neutralizing activity was found in sera from animals immunized with 4 log₁₀ FFU of the second component of YF d-PIV packaged in YF envelope. No YF-specific neutralizing response was induced by two doses of PIV-YF/WN packaged into YF envelope, and similarly, no WN-specific response was induced by two doses of PIV-JE packaged into WN envelope. The absence of neutralizing response against packaging envelopes permits manufacturing different PIV vaccines in one (universal) manufacturing helper cell line, or immunization of one individual with different recombinant vaccines based on the same vector, according to the present invention. PIV Cocktails

Because PIVs undergo a single (optionally several, but limited) round(s) of replication *in vivo*, we considered that mixtures of different PIV vaccines can be administered without interference between individual constructs in the mixture (cocktail). To elucidate whether PIV vaccines can be used in cocktail formulations, immune responses in mice and hamsters to several PIV constructs given as mixtures were compared to the same constructs given individually. Similar results were obtained in both animal models. Results of mouse experiments are shown in Table 6. Similar anti-JE neutralizing antibody titers were observed in pools of sera from animals that were given one or two doses of either PIV-JE+PIV-WN mixture or PIV-JE alone (1:20 vs. 1:80 and 1:640 vs. 1:160, for one and two doses, respectively). Similarly, WN-specific titers against PIV-JE+PIV-WN mixture and PIV-WN alone were similar (1:320 vs. 1:640 and 1:5,120 vs. 1:5,120 for one and 2 doses, respectively). No or little cross-specific response was induced by either PIV-JE or -WN. The result was also confirmed by measuring PRNT₅₀ titers in sera from individual animals. Thus, it is clear that PIV vaccines can be efficiently administered as cocktails, inducing immunity against two or more flavivirus pathogens. In addition, as discussed above, various cocktails can be made between non-flavivirus PIV vaccines, or between any of flavivirus and non-flavivirus PIV vaccines.

In Vitro Studies

Different PIV prototypes were serially passaged up to 10 times in helper BHK cells, for s-PIVs, or in regular Vero cells, for d-PIVs. Samples harvested after each passage were titrated in Vero cells by immunostaining. Constructs grew to high titers, and no recombination restoring whole virus was

observed. For instance, PIV-WN consistently grew to titers 7-8 \log_{10} FFU/ml in BHK-CprME(WN) helper cells (containing a VEE replicon expressing the WN virus C-prM-E proteins), and WN d-PIV grew to titers exceeding 8 \log_{10} FFU/ml in Vero cells, without recombination.

Example 2

PIV-TBE

PIV-TBE vaccine candidates can be assembled based entirely on sequences from wt TBE virus or the closely serologically related Langat (LGT) virus (naturally attenuated virus, e.g., wt strain TP-21 or its empirically attenuated variant, strain E5), or based on chimeric sequences containing the backbone (capsid and non-structural sequences) from YF 17D or other flaviviruses, such as WN virus, and the prM-E envelope protein genes from TBE, LGT, or other serologically related flaviviruses from the TBE serocomplex. YF/TBE LAV candidates are constructed based on the backbone from YF 17D and the prM-E genes from TBE or related viruses (e.g., the E5 strain of LGT), similar to other chimeric LAV vaccines.

Construction of PIV-TBE and YF/TBE LAV vaccine prototypes was performed by cloning of appropriate genetic elements into plasmids for PIV-WN (Mason et al., *Virology* 351:432-443, 2006; Suzuki et al., *J. Virol.* 82:6942-6951, 2008), or plasmids for chimeric LAVs (e.g., pBSA-AR1, a single-plasmid version of infectious clone of YF/JE LAV; WO 2008/036146), respectively, using standard methods in the art of reverse genetics. The prM-E sequences of TBE virus strain Hypr (GenBank accession number U39292) and LGT strain E5 (GenBank accession number AF253420) were first computer codon-optimized to conform to the preferential codon usage in the human genome, and to eliminate nucleotide sequence repeats longer than 8 nt to ensure high genetic stability of inserts (if determined to be necessary, further shortening of nt sequence repeats can be performed). The genes were chemically synthesized and cloned into plasmids for PIV-WN and YF/JE LAV, in place of corresponding prM-E genes. Resulting plasmids were in vitro transcribed and appropriate cells (Vero for chimeric viruses, and helper BHK cells for PIV) were transfected with RNA transcripts to generate virus/PIV samples.

YF/TBE LAV Constructs

In YF/TBE constructs containing either the TBE Hypr (plasmids p42, p45, and p59) or LGT E5 (plasmid P43) prM-E genes, two different types of the C/prM junction were first examined (see in FIG. 6; C/prM junctions only are shown in Sequence Appendix 1, and complete 5'-terminal sequences covering the 5'UTR-C-prM-E-beginning of NS1 region are shown in Sequence Appendix 2). The p42-derived YF17D/Hypr chimera contained a hybrid YF17D/Hypr signal peptide for the prM protein, while the p45-derived YF17D/Hypr chimera contained a hybrid YF17D/WN signal peptide for prM (Sequence Appendix 1). The former chimeric virus produced very high titers at both P0 (immediately after transfection) and P1 (the next passage in Vero cells), up to 7.9 \log_{10} PFU/ml, which were 0.5 \log_{10} times higher, compared to the latter virus; in addition it formed significantly larger plaques in Vero cells (FIG. 6). Thus, use of TBE-specific residues in the signal peptide for prM conferred a significant growth advantage over the signal containing WN-specific residues. The p43-derived YF17D/LGT chimera had the same prM signal as the p42-derived virus; it also produced very high titers at P0 and P1 passages (up to 8.1 \log_{10} PFU/ml) and formed large plaques. A derivative of the p42-derived virus was also pro-

duced from plasmid p59, which contained a strong attenuating mutation characterized previously in the context of a YF/WN LAV vaccine virus, specifically, a 3-a.a. deletion in the YF17D-specific C protein (PSR, residues 40-42 in the beginning of α -Helix I; WO 2006/116182). As expected, the p59 virus grew to lower titers (5.6 and 6.5 \log_{10} PFU/ml at P0 and P1, respectively), and formed small plaques (determined in a separate titration experiment and thus not shown in FIG. 6), compared to the parent p42-derived chimera. These initial observations of growth properties of YF/TBE LAV prototypes, and correlation of replication in vitro with plaque morphologies, have been confirmed in growth curve experiments (FIG. 8).

PIV-TBE Constructs

PIV-WN/TBE variants were constructed, and packaged PIV samples were derived from plasmids p39 and p40 (FIG. 7; Sequence Appendix 1 for C/prM junction sequences, and Sequence Appendix 3 for complete 5'UTR- Δ C-prM-E-beginning of NS1 sequences). These contained complete Hypr or WN prM signals, respectively. Both PIVs were successfully recovered and propagated in BHK-CprME(WN) or BHK-C(WN) helper cells (Mason et al., *Virology* 351:432-443, 2006; Widman et al., *Vaccine* 26:2762-2771, 2008). The P0 and P1 sample titers of the p39 variant were 0.2-1.0 \log_{10} times, higher than p40 variant. In addition, Vero cells infected with p39 variant were stained brighter in immunofluorescence assay using a polyclonal TBE-specific antibody, compared to p40, indicative of more efficient replication (FIG. 7). The higher rate of replication of the p39 candidate than p40 candidate was confirmed in a growth curve experiment (FIG. 8). In the latter experiment, both candidates appeared to grow better in the BHK-C(WN) helper cells compared to BHK-CprME(WN), with the p39 variant reaching titer of $\sim 7 \log_{10}$ PFU/ml on day 5 (note that peak titers have not been reached). The discovery of the effect of prM signal on replication rates of both PIV and chimeric LAV vaccine candidates, and head-to-head comparison of different signals to generate the most efficiently replicating and immunogenic (see above) construct, are a distinguishing feature of our approach. As discussed above, the invention also includes the use of other flavivirus signals, including with appropriate mutations, the uncoupling the viral protease and signalase cleavages at the C/prM junction, e.g., by mutating or deleting the viral protease cleavage site at the C-terminus of C preceding the prM signal, the use of strong non-flavivirus signals (e.g., tPA signal, etc.) in place of prM signal, as well as optimization of sequences downstream from the signalase cleavage site.

Other PIV-TBE variants based entirely on wt TBE (Hypr strain) and LGT virus (TP21 wild type strain or attenuated E5 strain), and chimeric YF 17D backbone/prM-E (TBE or LGT) sequences are also included in the invention. Helper cells providing appropriate C, C-prM-E, etc., proteins (e.g., TBE-specific) for trans-complementation can be constructed by means of stable DNA transfection or through the use of an appropriate vector, e.g., an alphavirus replicon, such as based on VEE strain TC-83, with antibiotic selection of replicon-containing cells. Vero and BHK21 cells can be used in practice of the invention. The former are an approved substrate for human vaccine manufacture; any other cell line acceptable for human and/or veterinary vaccine manufacturing can be also used. In addition to s-PIV constructs, d-PIV constructs can also be assembled. To additionally ascertain safety for vaccinees and the environment, appropriate modifications can be employed, including the use of degenerate codons and complementary mutations in the 5' and 3' CS elements, to minimize chances of recombination that theoretically could result in viable virus. Following construction, all vaccine

candidates can be evaluated in vitro for manufacturability/stability, and in vivo for attenuation and immunogenicity/efficacy, in available pre-clinical animal models, such as those used in development and quality control of TBE and YF vaccines.

Neurovirulence and Neuroinvasiveness in Mice of PIV-TBE and YF/TBE LAV Constructs

Young adult ICR mice (~3.5 week-old), were inoculated with graded doses of PIV-TBE and YF/TBE LAV candidates by the IC route to measure neurovirulence, or IP route to measure neuroinvasiveness (and later immunogenicity/efficacy). Animals that received $5 \log_{10}$ FFU of PIV-Hypr (p39 and p40) variants by both routes survived and showed no signs of sickness, similar to mock-inoculated animals (Table 7), and thus PIV-TBE vaccines are completely avirulent. Mice inoculated IC with YF 17D control ($1-3 \log_{10}$ PFU) showed dose-dependent mortality, while all animals inoculated IP ($5 \log_{10}$ PFU) survived, in accord with the knowledge that YF 17D virus is not neuroinvasive. All animals that received graded IC doses ($2-4 \log_{10}$ PFU) of YF/TBE LAV prototypes p42, p45, p43, and p59 died (moribund animals were humanely euthanized). These variants appear to be less attenuated than YF 17D, e.g., as evidenced by complete mortality and shorter AST at the $2 \log_{10}$ PFU dose, the lowest dose tested for YF/TBE LAV candidates. The non-neurovirulent phenotype of PIV-TBE, virulent phenotype of YF/TBE LAV and intermediate-virulence phenotype of YF 17D are also illustrated in FIG. 9, showing survival curves of mice after IC inoculation. It should be noted that the p43 (LGT prM-E genes) and p59 (the dC2 deletion variant of YF/Hypr LAV) were less neurovirulent than p42 and p45 YF/Hypr LAV constructs as evidenced by larger AST values for corresponding doses (Table 7). In addition, p43 and p59 candidates were non-neuroinvasive, while p42 and p45 caused partial mortality after IP inoculation ($5 \log_{10}$ PFU/dose) (Table 7; FIG. 10). It should be noted however that all the YF/TBE LAV constructs were significantly attenuated as compared to wt TBE viruses, e.g., compared to wt TBE Hypr virus, which is uniformly highly virulent for mice, both at very low IC ($LD_{50} \sim 0.1$ PFU) and IP ($LD_{50} \leq 10$ PFU) doses (Wallner et al., *J. Gen. Virol.* 77:1035-1042, 1996; Mandl et al., *J. Virol.* 72:2132-2140, 1998; Mandl et al., *J. Gen. Virol.* 78:1049-1057, 1997).

Immunogenicity/Efficacy of PIV-TBE and YF/TBE LAV Constructs in Mice

TBE-specific neutralizing antibody responses in mice immunized IP with one or two doses of the PIV-TBE or YF/TBE LAV variants described above, or a human formalin-inactivated TBE vaccine control (1:30 of human dose) are being measured. Animals have been challenged with a high IP dose (500 PFU) of wt Hypr TBE virus; morbidity (e.g., weight loss), and mortality after challenge are monitored. Immunogenicity/Efficacy of PIV-TBE and YF/TBE LAV Constructs in Mice

TBE-specific neutralizing antibody responses in mice immunized IP with one or two doses of the PIV-TBE or YF/TBE LAV variants described above (from experiment in Table 7), or a human formalin-inactivated TBE vaccine control (1:20 of human dose; one or two doses), or YF 17D and mock controls, were measured on day 20 by PRNT₅₀ against wt TBE Hypr virus (Table 8; second dose of indicated test

articles was given on day 14). [Titers were determined in individual sera, or pooled sera from two animals in most cases, or pooled sera from 4 animals for the YF17D and Mock negative controls]. Titers in individual test samples as well as GMTs for each group are provided in Table 8. Titers in test samples were similar within each group, e.g., in groups immunized with PIVs, indicating high uniformity of immune response in animals. As expected, no TBE-specific neutralizing antibodies were detected in negative control groups (YF 17D and Mock; GMTs < 1:10); accordingly, animals in these groups were not protected from challenge on day 21 post-immunization with a high IP dose (500 PFU) of wt Hypr TBE virus. Mortalities from partial observation (on day 9 post-challenge; observation being continued) are provided in Table 8, and dynamics of average post-challenge body weights indicative of morbidity are shown in FIG. 11. Neutralizing antibodies were detected in killed vaccine controls, which were particularly high after two doses (GMT 1:1,496); animals in the 2-dose group were completely protected in that there was no mortality or body weight loss (but not animals in the 1-dose group). Animals that received both one and two doses of PIV-Hypr p39 had very high antibody titers (GMTs 1:665 and 1:10,584) and were solidly protected, demonstrating that robust protective immunity can be induced by s-PIV-TBE defective vaccine. The two animals that survived immunization with YF/Hypr p42 chimera (see in Table 7) also had high antibody titers (GMT 1:6,085) and were protected (Table 8; FIG. 11). Interestingly, PIV-Hypr p40 and YF/Hypr p45 were poorly immunogenic (GMTs 1:15 and 1:153 for one and two doses, and 1:68, respectively). As discussed above, these contained WN-specific sequences in the signal for prM, while the highly immunogenic PIV-Hypr p39 and YF/Hypr p42 constructs contained TBE-specific signal sequences. In agreement with discussion above, this result demonstrates the importance of choosing the right prM signal, e.g., the TBE-specific signal, to achieve high-level replication/VLP secretion, which in this experiment in vivo resulted in drastically different immune responses. Immunogenicity of YF/LGT p43 and YF/Hypr dC2 p59 chimeras was relatively low which could be expected, because of the use of a heterologous envelope (LGT, different from challenge TBE virus) and high attenuating effect of the dC2 deletion, respectively.

Example 3

Foreign Gene Expression

In the examples of recombinant PIV constructs described below, genes of interest were codon optimized (e.g., for efficient expression in a target vaccination host) and to eliminate long nt sequence repeats to increase insert stability (≥ 8 nt long; additional shortening of repeats can be performed if necessary), and then chemically synthesized. The genes were cloned into PIV-WN vector plasmids using standard methods of molecular biology well known in the art, and packaged PIVs were recovered following in vitro transcription and transfection of appropriate helper (for s-PIVs) or regular (for d-PIVs) cells.

Expression of Rabies Virus G Protein in WN s-PIV and d-PIV Rabies virus, Rhabdoviridae family, is a significant human and veterinary pathogen. Despite the availability of several

(killed) vaccines, improved vaccines are still needed for both veterinary and human use (e.g. as an inexpensive pre-exposure prophylactic vaccines). Rabies virus glycoprotein G mediates entry of the virus into cells and is the main immunogen. It has been expressed in other vectors with the purpose of developing veterinary vaccines (e.g., Pastoret and Brochier, *Epidemiol. Infect.* 116:235-240, 1996; Li et al., *Virology* 356:147-154, 2006).

Full length rabies virus G protein (original Pasteur virus isolate, GenBank accession number NC_001542) was codon-optimized, chemically synthesized, and inserted adjacent to the ΔC , $\Delta prM-E$ and $\Delta C-prM-E$ deletions in PIV-WN vectors (FIG. 12). The sequences of constructs are provided in Sequence Appendix 4. General designs of the constructs are illustrated in FIG. 13. The entire G protein containing its own signal peptide was inserted in-frame downstream from the WN C protein either with the ΔC deletion (ΔC and $\Delta C-prM-E$ constructs) or without ($\Delta prM-E$) and a few residues from the prM signal. Foot and mouth disease virus (FMDV) 2A autoprotease was placed downstream from the transmembrane C-terminal anchor of G to provide cleavage of C-terminus of G from the viral polyprotein during translation. The FMDV 2A element is followed by WN-specific signal for prM and prM-E-NS1-5 genes in the ΔC construct, or signal for NS1 and NS1-5 genes in $\Delta prM-E$ and $\Delta C-prM-E$ constructs.

Packaged WN(ΔC)-rabiesG, WN($\Delta prM-E$)-rabiesG, and WN($\Delta C-prM-E$)-rabiesG PIVs were produced by transfection of helper BHK cells complementing the PIV vector deletion [containing a Venezuelan equine encephalitis virus (strain TC-83) replicon expressing WN virus structural proteins for trans-complementation]. Efficient replication and expression of rabies G protein was demonstrated for the three constructs by transfection/infection of BHK-C(WN) and/or BHK-C-prM-E(WN) helper cells, as well as regular BHK cells, by immunostaining and immunofluorescence assay (IFA) using anti-Rabies G monoclonal antibody (RabG-Mab) (FIG. 14). Titers were determined in Vero cells by immunostaining with the Mab or an anti-WN virus polyclonal antibody. Growth curves of the constructs in BHK-C-prM-E(WN) cells after transfection with in vitro RNA transcripts are shown in FIG. 14, bottom panels. The PIVs grew efficiently to titers ~6 to >7 log₁₀ FFU/ml. Importantly, nearly identical titers were detected by both RabG-Mab and WN-antibody staining, which was the first evidence of genetic stability of the insert. In PIV-infected Vero cells, which were fixed but not permeabilized, strong membrane staining was observed by RabG-Mab staining, demonstrating that the product was efficiently delivered to the cell surface (FIG. 15). The latter is known to be the main prerequisite for high immunogenicity of expressed G. Individual packaged PIVs can spread following infection of helper BHK cells, but cannot spread in regular cells as illustrated for WN(ΔC)-rabiesG PIV in FIG. 16. The fact that there is no spread in naïve BHK cells demonstrates that the recombinant RNA genomes cannot be non-specifically packaged into membrane vesicles containing the G protein, if produced by PIV infected cells. An identical result was obtained with the G protein of another rhabdovirus, Vesicular stomatitis virus (VSV), contrary to previous observations of non-specific packaging of Semliki Forest virus (SFV) replicon expressing VSV G protein (Rolls et al., *Cell* 79:497-506,

1994). The latter is a desired safety feature. [Alternatively, some non-specific packaging could result in a limited spread of PIV in vivo, potentially enhancing anti-rabies immune response. The latter could be also a beneficial feature, given that such PIV is demonstrated to be safe]. The stability of the rabies G insert in the three PIVs was demonstrated by serial passages in helper BHK-CprME(WN) cells at high or low MOI (0.1 or 0.001 FFU/cell). At each passage, cell supernatants were harvested and titrated in regular cells (e.g., Vero cells) using immunostaining with an anti-WN polyclonal antibody to determine total PIV titer, or anti-rabies G monoclonal antibody to determine titer of particles containing the G gene (illustrated for MOI 0.1 in FIG. 17; similar results were obtained at MOI 0.001). The WN(ΔC)-rabiesG PIV was stable for 5 passages, while the titer of insert-containing PIV started declining at passage 6, indicative of insert instability. This could be expected, because in this construct, large G gene insert (~1500 nt) is combined with a small ΔC deletion (~200 nt), significantly increasing the overall size of the recombinant RNA genome. In contrast, in WN($\Delta prM-E$)-rabiesG, and WN($\Delta C-prM-E$)-rabiesG PIVs, the insert is combined with a much larger deletion (~2000 nt). Therefore, these constructs stably maintained the insert for all 10 passages examined (FIG. 17). Further, it can be seen in FIG. 17 that at some passages, titers as high as 8 log₁₀ FFU/ml, or higher, were attained for all three PIVs, additionally demonstrating that PIVs can be easily propagated to high yields.

Following inoculation in vivo individually, the WN(ΔC)-rabiesG s-PIV is expected to induce strong neutralizing antibody immune responses against both rabies and WN viruses, as well as T-cell responses. The WN($\Delta prM-E$)-rabiesG and WN($\Delta C-prM-E$)-rabiesG PIVs will induce humoral immune response only against rabies because they do not encode the WN prM-E genes. WN(ΔC)-rabiesG s-PIV construct can be also co-inoculated with WN($\Delta prM-E$)-rabiesG construct in a d-PIV formulation (see in FIG. 12), increasing the dose of expressed G protein, and with enhanced immunity against both pathogens due to limited spread. As an example of spread, titration results in Vero cells of a s-PIV sample, WN($\Delta prM-E$)-rabiesG, and a d-PIV sample, WN($\Delta prM-E$)-rabiesG+WN(ΔC) PIV (the latter did not encode rabies G protein), are shown in FIG. 18. Infection of naïve Vero cells with s-PIV gave only individual cells stainable with RabG-Mab (or small clusters formed due to division of cells). In contrast, large foci were observed following infection with the d-PIV sample (FIG. 18, right panel) that were products of coinfection with the two PIV types.

The WN($\Delta C-prM-E$)-rabiesG construct can be also used in a d-PIV formulation, if it is co-inoculated with a helper genome providing C-prM-E in trans (see in FIG. 12). For example it can be a WN virus genome containing a deletion of one of the NS proteins, e.g., NS1, NS3, or NS5, which are known to be trans-complementable (Khromykh et al., *J. Virol.* 73:10272-10280, 1999; Khromykh et al., *J. Virol.* 74:3253-3263, 2000). We have constructed a WN-ANS 1 genome (sequence provided in Sequence Appendix 4) and obtained evidence of co-infection with WN($\Delta prM-E$)-rabiesG or WN($\Delta C-prM-E$)-rabiesG constructs, and spread in vitro, by immunostaining. In the case of such d-PIVs, rabies G protein can be also inserted and expressed in helper genome, e.g., WN- $\Delta NS1$ genome, to increase the amount of expressed rabies G protein resulting in an increased anti-rabies immune response. As

with any d-PIV versions, one immunogen can be from one pathogen (e.g., rabies G) and the other from a second pathogen, resulting in three antigenic specificities of vaccine. As

rabies virus G protein. The modification is intended to elucidate whether the use of a heterologous signal can increase the rate of F protein synthesis and/or replication of PIVs.

TABLE 1

| PIV prototype constructs used in platform development studies | | |
|---|--|---|
| Construct | Genetic composition | Packaged in |
| PIV-WN | wt NY99 WN virus | WN envelope; BHK-CprME(WN) or BHK-C(WN) helper cells (Mason et al., <i>Virology</i> 2006, 351: 432-43; Widman et al., <i>Vaccine</i> 2008, 26: 2762-71) |
| PIV-YF/WN | Envelope (VLP): wt WN NY99 Backbone: YF 17D | YF 17D envelope; BHK-CprME(YF) helper cells (Widman et al., <i>Adv Virus Res.</i> 2008, 72: 77-126) |
| PIV-WN/JE | Envelope (VLP): wt JE Nakayama Backbone: wt WN NY99 | JE or WN envelope; BHK-C(WN) or BI-1K-CprME(WN) helper cells (Ishikawa et al., <i>Vaccine</i> 2008, 26: 2772-8) |
| PIV-YF | YF 17D | YF 17D envelope; BHK-CprME(YF) or BHK-C(YF) helper cells (Mason et al., <i>Virology</i> 2006, 351: 432-43) |

discussed above, Δ NS1 deletions can be replaced with or used in combination with Δ NS3 and/or Δ NS5 deletions/mutations, in other examples.

Expression of RSV F Protein in WN s-PIV and d-PIV

Respiratory syncytial virus (RSV), member of Paramyxoviridae family, is the leading cause of severe respiratory tract disease in young children worldwide (Collins and Crowe, *Respiratory Syncytial Virus and Metapneumovirus*, In: Knipe et al. Eds., *Fields Virology*, 5th ed., Philadelphia: Wolters Kluwer/Lippincott Williams and Wilkins, 2007:1601-1646). Fusion protein F of the virus is a lead viral antigen for developing a safe and effective vaccine. To avoid post-vaccination exacerbation of RSV infection observed previously with a formalin-inactivated vaccine candidate, a balanced Th1/Th2 response to F is required which can be achieved by better TLR stimulation, a prerequisite for induction of high-affinity antibodies (Delgado et al., *Nat. Med.* 15:34-41, 2009), which should be achievable through delivering F in a robust virus-based vector. We have previously demonstrated the capacity of yellow fever virus-based chimeric LAV vectors to induce a strong, balanced Th1/Th2 response in vivo against an influenza antigen (WO 2008/036146). In the present invention, both yellow fever virus-based chimeric LAVs and PIV vectors are used for delivering RSV F to induce optimal immune response profile. Other LAVs and PIV vectors described herein can also be used for this purpose.

Full-length RSV F protein of A2 strain of the virus (GenBank accession number P03420) was codon optimized as described above, synthesized, and cloned into plasmids for PIV-WN s-PIV and d-PIV, using the insertion schemes shown in FIGS. 12 and 13 for rabies G protein, by applying standard methods of molecular biology. Exact sequences of the insertions and surrounding genetic elements are provided in Sequence Appendix 5. In vitro RNA transcripts of resulting WN(Δ C)—RSV F, WN(Δ prME)-RSV F, and WN(Δ CprME)-RSV F PIV constructs were used to transfect helper BHK-CprME(WN) cells. Efficient replication and expression of RSV F protein was first demonstrated by immunostaining of transfected cells with an anti-RSV F Mab, as illustrated for the WN(Δ prME)-RSV F construct in FIG. 19. The presence of packaged PIVs in the supernatants from transfected cells (titer as high as 7 log₁₀ FFU/ml) was determined by titration in Vero cells with immunostaining. Additionally, similar constructs can be used that contain a modified full length F protein gene. Specifically, the N-terminal native signal peptide of F is replaced in modified F protein with the one from

TABLE 2

| Safety: Suckling mouse neurovirulence ¹ | | | |
|--|----------------------------|---------------|-------------------------|
| Construct | Doses (log ₁₀) | Mortality (%) | AST (days) ² |
| PIV-YF | 1-4 | 0/10 (0%) | na |
| PIV-WN | 2-5 | 0/10 (0%) | na |
| PIV-WN/JE | 1-4 | 0/11 (0%) | na |
| PIV-YF/WN | 1-4 | 0/10-11 (0%) | na |
| WN d-PIV | 1-4 | 0/10-11 (0%) | na |
| YF d-PIV | 1-4 | 0/10 (0%) | na |
| YF17D | 2 | 10/10 (100%) | 7.6 |
| | 1 | 10/10 (100%) | 9.3 |
| | 0 | 9/10 (90%) | 9.9 |
| | -1 | 3/10 (30%) | 9.6 |
| YF/JE | 4 | 9/11 (82%) | 9.7 |
| | 3 | 7/10 (70%) | 12.3 |
| | 2 | 3/11 (27%) | 12 |
| | 1 | 0/11 (0%) | na |
| YF/WN | 3 | 2/11 (18%) | 12.5 |
| | 0-2 | 0/10-11 (0%) | na |

¹Single dose, IC inoculation, ICR 5-day old mice, graded log doses administered.

²AST for mice that died; na, not applicable.

TABLE 3

| PIV highly immunogenic and efficacious in mice ¹ | | | | |
|---|-----------------------------------|-------------|-------------|------------------------------|
| Group | Dose | PRNT Day 20 | PRNT Day 34 | Post-challenge mortality (%) |
| PIV-WN | 10 ⁵ | 640 | 1280 | 0/8 (0%) |
| | 10 ⁶ | 1280 | 2560 | 1/8 (12.5%) |
| | 10 ⁶ + 10 ⁵ | 2560 | 2560 | 0/6 (0%) |
| YF/WN control | 10 ⁴ | 1280 | 2560 | 1/8 (12.5%) |
| PIV-WN/JE | 10 ⁴ | 10 | 20 | N/D |
| | 10 ⁵ | 20 | 20 | N/D |
| | 10 ⁵ + 10 ⁵ | 20 | 160 | N/D |
| YF/JE control | 10 ⁴ | 160 | 320 | N/D |
| PIV-YF | 10 ⁴ | <10 | <10 | 8/8 (100%) |
| | 10 ⁵ | <10 | <10 | 5/7 (71%) |
| | 10 ⁵ + 10 ⁵ | 10 | 10 | 2/5 (40%) |
| YF17D control | 10 ⁴ | 640 | 1280 | 0/7 (0%) |
| Mock - WN challenge | Diluent | N/D | 0 | 7/7 (100%) |
| control - YF challenge | Diluent | N/D | 0 | 8/8 (100%) |

¹IP immunization (d0 prime, and d21 boost in select groups); challenge on d35: wt WN NY99, 3 log₁₀ PFU IP, 270 LD50; wt YF Asibi, 3 log₁₀ PFU IC, 500 LD50; N/D, not determined.

TABLE 4

| PIV are immunogenic in hamsters and protect against challenge ¹ | | | | | |
|--|-----------------------------------|----------------|-----------|------------|--------------------|
| Group | Dose(s) | POST-CHALLENGE | | | |
| | | PRNT Day 38 | Mortality | Morbidity | Peak viremia (log) |
| PIV-WN | 10 ⁵ | 320 | 0/5 (0%) | 0/5 (0%) | 2.3 |
| | 10 ⁶ | 640 | 0/5 (0%) | 0/5 (0%) | 1.8 |
| | 10 ⁶ + 10 ⁵ | 1280 | 0/5 (0%) | 0/5 (0%) | <1.3 |
| YF/WN control | 10 ⁴ | ≥2560 | 0/5 (0%) | 0/5 (0%) | <1.3 |
| PIV-WN/JE | 10 ⁴ | 20 | 2/5 (40%) | 2/5 (40%) | 2.2 |
| | 10 ⁵ + 10 ⁵ | 40 | 0/5 (0%) | 0/5 (0%) | <1.3 |
| YF/JE control | 10 ⁴ | 2560 | 0/5 (0%) | 0/5 (0%) | 1.3 |
| PIV-YF | 10 ⁴ | <10 | 1/3 (33%) | 3/3 (100%) | 8.3 |
| | 10 ⁵ | <10 | 1/5 (20%) | 4/5 (80%) | 8.3 |
| | 10 ⁵ + 10 ⁵ | 20 | 0/4 (0%) | 0/4 (0%) | 2.5 |
| YF17D control | 10 ⁴ | ≥2560 | 0/4 (0%) | 0/4 (0%) | <1.3 |
| Mock control - WN challenge | Diluent | <10 | 3/4 (75%) | 4/4 (100%) | 4.0 |
| - YF challenge | Diluent | <10 | 1/4 (25%) | 4/4 (100%) | 8.4 |
| - JE challenge | Diluent | <10 | 2/5 (40%) | 2/5 (40%) | 3.0 |

¹Syrian hamsters, SC inoculation (d0, and d21 in select groups); challenge (d39): wt WN NY385/99 6 log₁₀ PFU IP, wt JE Nakayama 5.8 log₁₀ PFU IC, or hamster-adapted YF Asibi 7 log₁₀ PFU IP (McArthur et al., J. Virol. 77:1462-1468, 2003; McArthur et al., Virus Res. 110:65-71, 2005).

TABLE 5

| Immunization of hamsters with PIV: comparison of SC and IP routes | | | | | |
|---|----------------|-----|----------------------------|----------------|------|
| Inoculums | PRNT Day 20-21 | | Boost (log ₁₀) | PRNT Day 34-38 | |
| | SC | IP | | SC | IP |
| PIV-WN | 40 | 320 | 5 | 1280 | 1280 |
| PIV-YF/WN | 10 | 320 | 5 | 160 | 1280 |
| PIV-WN/JE | 10 | 80 | 5 | 40 | 640 |
| PIV-YF | <10 | 10 | 5 | 20 | 80 |

TABLE 6

| Immune responses to PIV cocktails (mice) ¹ | | | | | |
|---|-----------------------------------|-------------|---------|-------------|---------|
| Group | Dose | PRNT Day 20 | | PRNT Day 34 | |
| | | Anti-JE | Anti-WN | Anti-JE | Anti-WN |
| PIV-WN/JE + RV-WN | 10 ⁵ + 10 ⁵ | 20 | 320 | 640 | 5120 |
| PIV-WN/JE alone | 10 ⁵ | 80 | <10 | 160 | 20 |
| PIV-WN alone | 10 ⁵ | <10 | 640 | <10 | 5120 |
| Mock | — | <10 | <10 | <10 | <10 |

¹C57/BL6 mice, IP inoculations on days 0 and 21; pooled serum PRNT titers.

TABLE 7

| Neurovirulence (IC inoculation) and neuroinvasiveness (IP inoculation) of PIV-TBE and YF/TBE vaccine constructs in adult ICR mice | | | | | | |
|---|------------------------------|---------------|------------------------|------------------------------|---------------|------------------------|
| Con-struct | Neurovirulence (IC route) | | | Neuroinvasiveness (IP route) | | |
| | Dose(s) (log ₁₀) | Mortality (%) | AST, days ¹ | Dose(s) (log ₁₀) | Mortality (%) | AST, days ¹ |
| PIV-Hypr p39 | 5 | 0/7 (0%) | na | 5 | 0/16 (0%) | na |
| PIV-Hypr p40 | 5 | 0/6 (0%) | na | 5 | 0/16 (0%) | na |
| YF/Hypr p42 | 4 | 8/8 (100%) | 6.3 | 5 | 6/8 (75%) | 13.3 |
| YF/Hypr p43 | 3 | 8/8 (100%) | 6.4 | | | |
| YF/Hypr p44 | 2 | 8/8 (100%) | 7.4 | | | |
| YF/Hypr p45 | 4 | 8/8 (100%) | 7.9 | 5 | 0/8 (0%) | na |
| LGT | 3 | 8/8 (100%) | 7.6 | | | |
| YF/Hypr p45 | 2 | 8/8 (100%) | 8.4 | | | |
| YF/Hypr p45 | 4 | 8/8 (100%) | 6.1 | 5 | 5/8 (62.5%) | 112 |
| YF/Hypr p45 | 3 | 8/8 (100%) | 6.6 | | | |
| YF/Hypr p45 | 2 | 8/8 (100%) | 6.8 | | | |
| YF/Hypr p45 | 4 | 8/8 (100%) | 6.6 | 5 | 0/8 (0%) | na |
| YF/Hypr p45 | 3 | 8/8 (100%) | 7.4 | | | |
| dC2 p59 | 2 | 8/8 (100%) | 8.1 | | | |
| YF 17D | 3 | 8/8 (100%) | 9 | 5 | 0/8 (0%) | na |
| | 2 | 7/8 (87.5%) | 9.6 | | | |
| | 1 | 4/8 (50%) | 10 | | | |
| Mock (diluent) | none | 0/8 (0%) | na | none | 0/8 (0%) | na |

¹AST for mice that died.

TABLE 8

| Neutralizing antibody titers (PRNT ₅₀) in mice immunized IP (determined against wt TBE virus Hypr), and protection from challenge (postchallenge observation, day 9) | | | | |
|--|----------------------------|---|------------------------|---|
| Immunogen | Dose(s), log ₁₀ | PRNT ₅₀ titer, individ. samples ¹ | PRNT ₅₀ GMT | Postchallenge mortality (%) on day 9 ² |
| PIV-Hypr p39, 1 dose | 5 | 1746 (2) | 665 | 0/8 (0%) |
| | | 1187 (2) | | |
| | | 164 (2) | | |
| | | 574 (2) | | |

TABLE 8-continued

| Neutralizing antibody titers (PRNT ₅₀) in mice immunized IP (determined against wt TBE virus Hypr), and protection from challenge (postchallenge observation, day 9) | | | | |
|--|-------------------------------|--|------------------------|---|
| Immunogen | Dose(s), log ₁₀ | PRNT ₅₀ titer, individ. samples ¹ | PRNT ₅₀ GMT | Postchallenge mortality (%) on day 9 ² |
| PIV-Hypr p39, 2 doses | 5 + 5 | 16229 (2) 12928 (2) 12927 (2) 4627 (2) | 10,584 | 0/8 (0%) |
| PIV-Hypr p40, 1 dose | 5 | <10 (2) <10 (2) 18 (2) 33 (2) | 15 | 6/8 (75%) |
| PIV-Hypr p40, 2 doses | 5 + 5 | 169 (2) 638 (2) 26 (2) 192 (2) | 153 | 1/8 (12.5%) |
| YF/Hypr p42 | 5 | 9210 (1) 4020 (1) | 6,085 | 0/2 (0%) |
| YF/LGT p43 | 5 | 123 (2) 32 (2) 96 (2) 45 (2) | 64 | 1/8 (12.5%) |
| YF/Hypr p45 | 5 | 292 (2) 16 (1) | 68 | 0/3 (0%) |
| YF/Hypr dC2 p59 | 5 | 194 (2) 93 (2) 45 (2) 26 (2) | 68 | 0/8 (0%) |
| Killed human TBE vaccine, 1 dose (at 1/20 of human dose) | 1/20 | 19 (2) <10 (2) 13 (2) <10 (2) | 12 | 1/8 (12.5%) |
| Killed human TBE vaccine, 2 doses (each at 1/20 of human dose) | 1/20 + 1/20 | 3435 (2) 1267 (2) 770 (2) | 1,496 | 0/6 (0%) |
| YF 17D control | 5 | <10 (4) 11 (4) | <10 | 5/8 (62.5%) |
| Mock | none | <10 (4) <10 (4) | <10 | 4/8 (50%) |

¹Numbers in parenthesis correspond to number of mice in each pooled serum sample tested.

²Mortalities on day 9 are shown.

TABLE 9

| Examples of published attenuating E protein mutations that can be used for attenuation of chimeric TBE LAV candidates | | | | |
|---|--------|--|----------------|--|
| Residue | Domain | Comments | Attenuation in | Reference |
| N52R | II | DI-DII hinge, possibly involved in hinge motion required for fusion activation | JE, YF | Hasegawa et al, 1992, Schlesinger et al, 1996 |
| E84K | II | conserved, E in TBE, K/R in others, attenuated by passage in <i>Ixodes ricinus</i> ticks, DII contains flavivirus cross reactive epitopes | TBE | Labuda et al, 1994 |
| E85K | II | conserved, E in TBE, K/R in others, attenuation obtained as plaque variants in Vero cells, DII contains flavivirus cross reactive epitopes | JE | Wu et al, 1997 |
| H104K | II | within highly conserved fusion peptide (aa 98-113), H in TBE, G in others | TBE | Rey et al, 1995 |
| L107F | II | within highly conserved fusion peptide (aa 98-113), L in all flaviviruses, F in attenuated JE | TBE, JE, WN | Rey et al, 1995, Arroyo et al, 1999, 2004 |
| T123K | II | DI-DII hinge, T in TBE, A in KFD | TBE | Holzmann et al, 1997 |
| K126E | II | DI-DII hinge, K in TBE, E in D-2 | DEN2 | Bray, 98 |
| K136E | II | DI-DII hinge, K in TBE and JE, E in D-2 | JE | |
| N154L(Y) | I | glycosylation site, packed with conserved H 104, involved in fusion. | DEN2, DEN4, YF | Guirakhoo et al, 1993, Pletnev et al, 1993, Kawano et al, 1993, Jennings et al, 1994 |
| K171E | I | external edge of DI, involved in fusion | TBE | Mandl, 1989, Holzmann, 1997 |
| I173T | I | external edge of DI, involved in fusion | YF | Chambers and Nickells 2001 |
| D181Y | I | DI- DII hinge | TBE | Holzmann et al, 1997 |

TABLE 9-continued

| Examples of published attenuating E protein mutations that can be used for attenuation of chimeric TBE LAV candidates | | | | |
|---|--------|---|----------------|--|
| Residue | Domain | Comments | Attenuation in | Reference |
| K204R | | Lining Hydrophobic pocket, involve in fusion | DEN1, DEN3 | Guirakhoo et al, 2004 |
| P272S | II | highly conserved, junction of one the of 2 alpha helices | JE | Cecilia et al, 1991 |
| G308N | III | cell attachment, DKT in TBE, EGS in KFD, LI T-X in others, change to N produced glycosylation site in LI and reduced virulence, N-X-T/S glycosylation motif | | Jiang et al, 1993, Gao et al, 1994 |
| S310K | III | putative cell attachment, change from E to G in JE reduced virulence | JE | Jiang et al, 1993, Gao et al, 1994, Wu et al, 1997 |
| K311E | III | highly conserved, putative cell attachment | TBE, YF | Rey et al, 1995, Jennings, 1994 |
| T333L | III | putative cell attachment | YF, LGT | Raynman et al, 1998 |
| G334K | III | putative cell attachment | YF | Chambers and Nickells, 2001 |
| S335K | III | putative cell attachment | JE | Wu et al, 1997 |
| K336D | III | putative cell attachment | JE | Cecilia and Gould, 1991 |
| P337D | III | putative cell attachment | JE | Cecilia and Gould, 1991 |
| G368R | III | putative cell attachment | TBE, JE | Holzman et al 1997, Hasegawa et al 1992 |
| Y384H | III | change to H attenuated TBE, putative cell attachment, -3 position to deleted RGD in TBE | TBE | Holzmann et al, 1990 |
| V385R | III | conserved, -2 position to deleted RGD in TBE, putative cell attachment | D2 | Hiramatsu et al, 1996, Lobigs, 1990 |
| G386R | III | highly conserved, -1 position to deleted RGD in TBE, putative cell attachment | D2, MVE | Hiramatsu, 1996, Lobigs et al, 1990 |
| E387R | III | conserved, +2 position to deleted RGD in TBE, putative cell attachment | D2, MVE | Hiramatsu, 1996, Lobigs et al, 1990 |
| F403K | none | highly conserved, C-terminal region not included in crystal structure sE | D-2, D-4 | Kawano et al, 1993, Bray et al, 1998 |
| H438Y | None | highly conserved, C-terminal region not included in crystal structure sE | LGT | Campbell and Pletnev 2000 |
| H496R | none | highly conserved, C-terminal region not included in crystal structure sE | TBE | Gritsun et al, 2001 |

References: Hasegawa et al., *Virology* 191(1): 158-165; Schlesinger et al., *J. Gen. Virol.* 1996, 77 (Pt 6): 1277-1285, 1996; Labuda et al., *Virus Res.* 31(3): 305-315, 1994; Wu et al., *Virus Res.* 51(2): 173-181, 1997; Holzmann et al., *J. Gen. Virol.* 78 (Pt 1): 31-37, 1997; Bray et al., *J. Virol.* 72(2): 1647-1651, 1998; Guirakhoo et al., *Virology* 194(1): 219-223, 1993; Pletnev et al., *J. Virol.* 67(8): 4956-4963, 1993; Kawano et al., *J. Virol.* 67(11): 6567-6575, 1993; Jennings et al., *J. Infect. Dis.* 169(3): 512-518, 1994; Mandl et al., *J. Virol.* 63(2): 564-571, 1989; Chambers et al., *J. Virol.* 75(22): 10912-10922, 2001; Cecilia et al., *Virology* 181(1): 70-77, 1991; Jiang et al., *J. Gen. Virol.* 74 (Pt 5): 931-935, 1993; Gao et al., *J. Gen. Virol.* 75 (Pt 3): 609-614, 1994; Holzmann et al., *J. Virol.* 64(10): 5156-5159, 1990; Hiramatsu et al., *Virology* 224(2): 437-445, 1996; Lobigs et al., *Virology* 176(2): 587-595, 1990; Campbell et al., *Virology* 269(1): 225-237, 2000; Gritsun et al., *J. Gen. Virol.* 82 (Pt 7): 1667-1675, 2001.

Example 4

Delivery of SIV Gag and Env Proteins (HIV Prototypes) in WN s-PIV and d-PIV

An artificial cassette containing SIV (GenBank accession number ADM52218.1) gp120 (a modified gene where the native signal sequence was replaced with the tPA signal and gp41 was truncated to contain only the TM domain), Gag, and Pro (protease) genes is shown in FIG. 20. The cassette was designed in a way that would allow its expression in the recombinant PIV ORF as a single precursor (different from SIV or HIV gene organization). To allow for cleavage into individual SIV proteins, the genes are separated by FMDV 2A autoprotease sequences (see above). The nucleotide sequence of the entire cassette (~4 kb in length) was optimized by silent nucleotide changes to eliminate direct sequence repeats (e.g., all repeats longer than 8 nt were eliminated) to increase insert stability (using optimization algorithms at DNA 2.0) and by incorporating monkey codon preference to enable efficient protein translation in primate cells.

The codon-optimized cassette was chemically synthesized, followed by in-frame insertion of the genes, alone or in different combinations, in PIV-WN vectors in place of the ΔC (RV909 vector), AprM-E (RV230 vector) or ΔC-prM-E (dC RV230 vector) deletions. Examples of sequences of the constructs are provided in Sequence Appendix 6. Inserts of the

first three constructs in FIG. 20, starting with the Env glycoprotein, were designed similarly to the PIV WV-rabies G described hereinabove (gp120 signal fused with a portion of the signal sequence for prM at the end of the C gene or downstream from ΔC deletion depending on vector), as is also additionally illustrated for individual Env constructs in FIG. 21. In addition, alternate dC RV230 Env constructs were generated, in which the tPA signal and/or the SIV Env TM region of the gp120 gene were replaced with rabies virus G protein-specific signal and/or anchor sequences (three bottom constructs in FIG. 21), to determine whether these heterologous rabies G-derived sequences will have a beneficial effect on gp120 presentation or recombinant PIV replication. Gag and Gag-Pro insertions were designed to start with and end with FMDV 2A autoprotease sequences, to free the N- and C-termini of the cytoplasmically synthesized Gag protein. They were cloned in place of the AprM-E or AC-prM-E deletions (FIGS. 20 and 22). The N-terminal FMDV 2A was positioned either downstream from the viral cleavage site in C, or downstream from additional 9 or 18 amino acids following the cleavage site (from the prM signal) in the RV230 and dC RV230 vectors (FIG. 22) in order to determine which fusion type is preferable for efficient cleavage of FMDV 2A preceding Gag, which theoretically can be important in terms of both transgene expression and PIV replication.

Correct processing of the polyprotein in recovered SIV Gag and SIV Gag/Pro PIVs grown in helper cells was confirmed by Western blot using anti-Gag antibodies (FIG. 23).

Constructs expressing Gag alone showed the correct individual p58 Gag band of ~58 kDa, and constructs that also included Pro also showed an additional band of p28 which is a product of Gag cleavage by Pro. Immunostaining of naïve Vero cells infected with the Gag PIVs (constructs shown in FIG. 24D), showed individual stained cells as expected from sPIV (FIGS. 24A-C).

Efficient replication in vivo is illustrated by growth curves of SIV Gag PIV variants after transfection of helper cells with in vitro synthesized PIV RNA (P0 passage) (FIGS. 25A-F). Some of the PIV variants grew efficiently to titers in excess of $7 \log_{10}$ FFU/ml, and nearly identical titers were detected by both anti-Gag and anti-WN antibody staining, which was the first evidence of genetic stability of the Gag insert. When SIV Gag PIV was propagated in naïve Vero cells as a two-component formulation (d-PIV, sometimes also designated as tc-PIV), together with PIV-WN helper with ΔC deletion (RV909), titers in the excess of $8 \log_{10}$ FFU/ml were observed (FIG. 26). These results confirm that this formulation does not require helper cells for production (the principle of dPIV is described above).

High insert stability is illustrated for one of the SW Gag PIV variants in FIG. 27. The stability of Gag was examined by ten serial passages of a RV230-Gag variant, containing Gag gene in place of large $\Delta prM-E$ deletion, in helper BHK-CprME(WN) cells at MOI 0.1 FFU/cell. At each passage, cell supernatants were harvested and titrated in regular Vero cells using immunostaining with an anti-WN antibody to determine total PIV titer, or an anti-SW Gag antibody to determine titer of particles containing the Gag gene. Similar WN and Gag titers were observed after all 10 passages and no significant progressive decline in Gag positive titers was observed, e.g., as compared to the WN(ΔC)-rabies G PIV expressing the G insert in place of the very short AC deletion (see above).

Viable PIV-(WN)-SIV Env variants (FIGS. 20 and 21) were also recovered in helper BHK cells transfected with in vitro RNA transcripts and efficient expression of gp120 was demonstrated by immunofluorescence (FIGS. 28A-D and FIG. 29). Interestingly, efficient intracellular expression of the original gp120 was observed in Vero cells infected with packaged dC230Env variant as determined by immunostaining using anti-SIV Env antibody after methanol fixation (FIG. 28D), but little gp120 was detected on the surface of the infected cells fixed by formalin (FIG. 28B), indicating inefficient transport of the translation product through the secretory pathway or cleavage of the TM domain away from the gp120 molecule. The dC230Env/RabG anchor construct (FIG. 21), in which the SIV Env TM domain was replaced with the TM anchor sequence from rabies virus G protein, not only provided efficient intracellular expression of gp120 (FIG. 28C), but also enabled its efficient cell surface delivery (FIG. 28A and FIG. 29). Better surface expression/secretion of Env variants should result in higher immunogenicity of vaccine candidates. Therefore, the results presented with these constructs confirm the beneficial effect of using heterologous TM and/or signal sequences to increase immunogenicity of HIV Env glycoproteins.

Examples of sequences of similar PIV-HIV vaccine designs, using HIV-1 Clade C gene sequences, are provided in Appendix 7.

These examples demonstrate the feasibility of robust delivery of SIV (HIV) glycoproteins (e.g., variants of Env) as well as cytoplasmic antigens (Gag, Pol, Nef and any other desired intracellular antigens), some of which can be secreted as SIV/HIV VLPs (e.g., Gag with or without Env), by PIV vaccine vectors.

In addition to gp120, other variants of the HIV Env immunogen, such as the full-length gp160, gp140, gp145, gp41, etc., with or without desired mutations, truncations, deletions, or insertions (e.g., of dominant CD4 T cell epitopes, etc., including of non-HIV origin) in expressed molecules increasing immunogenicity and/or breadth of immune response against the variable HIV genotypes/strains, can be expressed without changing the meaning of this invention. Examples of possible modifications of Env are discussed below.

The Envelope (Env) protein is one of the primary targets of the humoral immune response upon infection with HIV. However, the Env protein has a number of defenses which prevent an effective antibody response from being mounted. These defenses include high degree of sequence variability, protection of functionally important domains through the use of variable loops and quaternary interactions, and high levels of glycosylation to shield the underlying protein backbone. In order to overcome this researchers have attempted a number of methods to increase the potency and breadth of antibody responses to Env. These modifications begin with an alteration of the underlying protein backbone itself. Attempts to minimize the genetic distance between immunizing isolates and those seen in the wild have led to the use of centralized sequences (consensus and ancestral) as immunogens (Kothe et al., *Virology* 2007, 360:218-234; Liao et al., *Virology* 2006, 353: 268-282). Modifying specific glycosylations has also been attempted. In some instances, hyperglycosylation of Env to mask unwanted epitopes in order to focus the humoral response on neutralizing domains has been utilized (Selvarajah et al., *J. Virol.* 2005, 79-12148-12163). Others have attempted to eliminate specific glycans to increase the availability of critical domains and hence increase Env immunogenicity (Li et al., *J. Virol.* 2008, 82:638-651). Altering the total glycosylation of the Env protein with expression in different systems has also been investigated (Kong et al., *J. Mol. Biol.* 2010, 403:131-147). Outside of post translational modifications other groups have focused on manipulating Env variable loops as a means to increase immunogenicity. These modifications include shortening or deletion of variable loops (Ching and Stamatatos, *J. Virol.* 2010, 84:9932-9946; Yang et al., *J. Virol.* 2004, 78:4029-4036) as means to expose underlying domains. On the surface of virions, functional Env spikes exist as non-covalently linked trimers. However, these trimers are highly unstable making them difficult to use as immunogens. To overcome this hurdle attempts have been made to stabilize these trimers through mutagenesis (Beddows et al., *J. Virol.* 2005, 79:8812-8827) and introduction of heterologous trimerization domains (Yang et al., *J. Virol.* 2002, 76:4634-4642). Attempts have also been made to graft known epitopes recognized by mAbs to heterologous scaffolds (Phogat et al., *Virology* 2008, 373: 72-84; Zolla-Pazner et al., *J. Virol.* 2011, 85:9887-9898). Others have attempted to overcome the low immunogenicity of HIV Env by combining Env with immunostimulatory molecules in an effort to nonspecifically raise the immunogenicity of immunization (Melchers et al., *J. Virol.* 2011, published ahead of print, doi:10.1128/JVI.06259-11).

If necessary, these and/or any other modifications of Env or other expressed HIV immunogens leading to increased immunogenicity and/or breadth of humoral or cellular responses can be incorporated in the HIV antigenic moieties of PIV-HIV without changing the meaning of this invention.

Example 5

Delivery of HA Protein of Influenza H1N1 Virus
(Strain New Caledonia) in WN s-PIV (which
Optionally can be Used in d-PIV)

The full-length HA gene of Flu strain New Caledonia was cloned in place of Δ prM-E and Δ C-prM-E deletions of PIV-WN vectors in the same fashion as described for Rabies G, RSV F, and SIV Env (as described above; FIG. 30). Examples of sequences are provided in Sequence Appendix 8. The variants were viable, and grew to high titers immediately after RNA transfections of helper cells (FIGS. 31A-B and FIGS. 32A-D). Identical titers in the growth curves using immunostaining with anti-WN and anti-HA antibodies provided evidence of insert stability.

All variants efficiently expressed the HA protein both intracellularly (methanol fixation) and on the cell surface (formalin fixation) of infected Vero cells as shown by immuno-fluorescence (FIGS. 33A-F, 34A-B, 35A-H, 36A-D, and 37A-B). The latter is a known prerequisite for high HA immunogenicity. Importantly, the expressed HA was efficiently recognized by both antibodies against the HA stem and HA globular head, confirming correct, native protein conformation (FIGS. 37A-B).

Other flu antigens can be similarly delivered, such as NA, M2 (e.g., M2e), etc., or fragments thereof. With respect to HA, various modifications can be introduced, and modified antigens then expressed in PIV vaccine vectors, without changing the meaning of this invention.

The PIV-SIV and PIV-Flu vaccine candidates described in Examples 4 and 5 can be tested for immunogenicity and efficacy in animal models. Earlier in vivo data have demonstrated that PIV vaccines expressing transgenes are highly immunogenic in animals, as has been shown for PIV-RSV F (see, e.g., WO 2010/107847, incorporated herein by reference), and more recent experiments for PIV-Rabies G.

The following Sequence Appendices include the following sequences:

Appendix 1

Construct 1: Sequence of RepliVax-WN (Δ CprME)-SIV 9AA FMD Gag (partial). DNA disclosed as SEQ ID NO: 79 and protein disclosed as SEQ ID NO: 80.

Construct 2: Sequence of RepliVax-WN (Δ CprME)-SIV 9AA FMD Gag & Pr (partial). DNA disclosed as SEQ ID NO: 81 and protein disclosed as SEQ ID NO: 82.

Construct 3: Sequence of RepliVax-WN (Δ CprME)-SIV Anch Gag (partial). DNA disclosed as SEQ ID NO: 83 and protein disclosed as SEQ ID NO: 84.

Construct 4: Sequence of RepliVax-WN (Δ CprME)-SIV Anch Gag & Pro (partial). DNA disclosed as SEQ ID NO: 85 and protein disclosed as SEQ ID NO: 86.

Construct 5: Sequence of RepliVax-WN (Δ CprME)-SIV FMD2a Gag. DNA disclosed as SEQ ID NO: 87 and protein disclosed as SEQ ID NO: 88.

Construct 6: Sequence of RepliVax-WN (Δ CprME)-SIV fmd2A Gag & Pr (partial). DNA disclosed as SEQ ID NO: 89 and protein disclosed as SEQ ID NO: 90.

Construct 7: Sequence of RepliVax-WN (Δ CprME)-SIV Env (partial). DNA disclosed as SEQ ID NO: 91 and protein disclosed as SEQ ID NO: 92.

Construct 8: Sequence of RepliVax-WN (Δ CprME)-SIV Env No Transmembrane (partial). DNA disclosed as SEQ ID NO: 93 and protein disclosed as SEQ ID NO: 94.

Construct 9: Sequence of RepliVax-WN (Δ CprME)-SIV ENV Rab G Transmembrane (TM) (partial). DNA disclosed as SEQ ID NO: 95 and protein disclosed as SEQ ID NO: 96.

Construct 10: Sequence of RepliVax-WN (Δ CprME)-SIV Env RabG Chimera, Signal Sequence and Transmembrane (TM). DNA disclosed as SEQ ID NO: 97 and protein disclosed as SEQ ID NO: 98.

Construct 11: Sequence of RepliVax-WN (Δ C)-SIV Env (partial). DNA disclosed as SEQ ID NO: 99 and proteins disclosed as SEQ ID NOS 100-102, respectively, in order of appearance.

Appendix 2

Sequence of RepliVax-WN (Δ prME)-HIV Gag (partial). DNA disclosed as SEQ ID NO: 103 and protein disclosed as SEQ ID NO: 104.

Sequence of RepliVax-WN (Δ prME)-HIV Env Gp140 (partial). DNA disclosed as SEQ ID NO: 105 and protein disclosed as SEQ ID NO: 106.

Appendix 3

Construct 1: Sequence of RepliVax-WN (Δ prME)-HA New Caledonia (partial). DNA disclosed as SEQ ID NO: 107 and protein disclosed as SEQ ID NO: 108.

Construct 2: Sequence of RepliVax-WN (Δ CprME)-HA New Caledonia (partial). DNA disclosed as SEQ ID NO: 109 and protein disclosed as SEQ ID NO: 110.

Sequence Appendix 1

CV-TBEV Hypr or CV-LGT E5 with YFV/TBEV chimeric signal (p42, p59, and p43 constructs)

```

          YF17D partial signal
          ~~~~~
C protein YF17D
-----
401  R K R R S H D V L T V Q F L I L G M L G M T I A A T V R
      A G G A A C G C C G T T C C C A T G A T G T T C T G A C T G T G C A A T T C C T A A T T T T G G G C A T G C T G G G C A T G A C A N T C G C A G C T A C G G T T C G C
      T C C T T T G C G G C A A G G G T A C T A C A G A C T G A C A C G T T A A G G A T T A A A A C C C G T A C G A C C C G T A C T G T T A G C G T C G A T G C C A A C G G
          TBEV partial signal
          ~~~~~
Hypr or LGT E5 prM protein
-----

```

CV-TBEV Hypr with YFV/WNV chimeric signal (p45)

```

          YF 17D partial signal
          ~~~~~
C protein YF17D
-----
401  R K R R S H D V L T V Q F L I L G M L A C V G A A T V R
      A G G A A C G C C G T T C C C A T G A T G T T C T G A C T G T G C A A T T C C T A A T T T T G G G C A T G C T G G C T T G T C T G G A G C A G C T A C C G T G C G A
      T C C T T T G C G G C A A G G G T A C T A C A G A C T G A C A C G T T A A G G A T T A A A A C C C G T A C G A C C G A A C A C A G C C T C G T C G A T G C C A C G T
          WNV partial signal
          ~~~~~
Hypr prM protein
-----

```

RV-WNV/TBEV Hypr with TBEV signal (p39)

```

          TBEV signal
          ~~~~~
WNV C protein
-----
201  Q K K R G G T D W M S W L I V I G M L G M P I A A T V R
      C A A A G A A A C G G G G G G A A C A G A C T G G A T G A G C T G G C T G C T C G T A A T C C G C A T G C T G G C C A T G A C A A T C G C A G C T A C G G T T C G C
      G T T T C T T T G C C C C C C T T G T C T G A C C T A C T C G A C C G A C G A G C A T T A G C C G T A C G A C C C G T A C T G T T A G C G T C G A T G C C A A G C G
          Hypr prM protein
          ~~~~~

```

RV-WNV/TBEV Hypr with WNV signal (p40)

```

          WNV signal
          ~~~~~
WNV C protein
-----
201  Q K K R G G K T G I A V M I G M L A C V G A A T V R
      C A A A G A A A C G G G G G G A A A G A C A G G C A T A C C T G T G A T G A T A G G C A T C G T G C T T G T G T C G G A G C A G C T A C C G T G C G A
      G T T T C T T T G C G C C C C C T T C T G T C C G T A C T C G A C A C T A C T A T C C G T A C G A C G A A C A C A G C C T C G T A C G A T G C G A T G C C A G C C T
          Hypr prM protein
          ~~~~~

```

Sequence Appendix 2

CV-TBEV Hypr with YFV/TBEV chimeric signal (p42)

```

                    5' UTR
-----
1  ASTAAATCCT GTGTGCTAAT TGAGGTGCAT TGGTCTGCAA ATCGAGTTGC TAGGCARTAA ACACATTTGG ATTAATTTTA ATCGTTCSTT GAGCGATTAG
   TCATTTAGGA CACACAGATTA ACTCCACGTA ACCAGACGTT TAGCTCAACG ATCCGTATTAT TGTGTAAACC TAATTAARAT TAGCAAGCAA CTCGCTAATC
   5' UTR
-----
                    C protein
-----
101 CAGAGAACTG ACCAGAACAT GTCGTGGTCT AAAGCTCAGG GAAAAACCCT GGGCGTCAAT ATGGTACGAC GAGGAGTTCG CTCCTTGPCA AACAAAATAA
   GTCCTTTGAC TGGTCTTGTA CAGACCAGCA TTTGAGTCC CTTTTGGGA CCCGCAGTAA TACCATGCTG CTCCTCAAGC GAGGAACAGT TGTCTTTATT
   C protein
-----
   Q K T K Q I G N R P G P S R G V Q G F F F F L F N I L T G K K I
201 AACAAAAAAC AAAACAATTT GGAACAGAC CTGGACCTTC AAGAGTGTG CAAGGATTEA TCTTTTCTT TTTGTTCAAC ATTTTGACTG GAAAAAAGAT
   TGTTTTTTGG TTTTGTATT CTTTGTCTG GACCTGGAAG TTCTCCACAA GTTCTTAAT ACAAAAGAA AACCAAGTTG TAAACTGAC CTTTTTCTAA
   C protein
-----
   T A H L K R L W K M L D P R Q G L A V L R K V K R V V A S L M R G
301 CACAGCCAC CTAAGAGGT TGTGAAAAT GCTGGACCA ACACAAGGCT TGGCTGTCT AAGAAAGTC AAGAGACTCG TGGCCAGTTT GATCAGAGGA
   GTGTCGGGTG GATTTCTCCA ACACCTTTA CGACCTGGGT TGTGTCCGA ACCGACAAGA TTCCCTTCAG TTCCTCACC ACCGGTCAA CTACTCTCCT
   YF17D partial signal
-----
                    TBEV partial signal
-----
                    C protein
-----
401 L S S R K R R S H D V L T V Q F L I L G M L G M T I A A T V R K E R
   TTGCTCAA GAAAGCGCG TTCCTATGAT GTTCTGACTG TGCAATTCCT AATTTTGGGC ATGCTGGGCA TGACAATCGC AGCTACGGTT CGCAAGGAAA
   AACACGAGTT CCTTTGCGGC AAGGTACTA CAAGACTGAC ACGTTAAGGA TTAAAACCCG TACGACCCGT ACIGTTAGCG TCGATGCCAA GCGTTCCTTT
   prM protein
-----
   D G S T V I R A E G K D A A T Q V R V E N G T C V I L A T D M G S
501 GAGACGGCAG TACGGTCATA CGCGCGGAAG GTAAGGATGC CGCTACCCAA GTGAGAGTGG AAAATGGTAC CTGCGTCATT CTGGCCACCG ACATGGGCTC
   CTCGCGGTC ATGCCAGTAT GCGCGCCTTC CATTCTACG GCGATGGGTI CACTCTCACC TTTTACCATG GACGCAATA GACCGGTGGC TGTACCCGAG
   prM protein
-----
   W C D D S L S Y E C V T I D Q G E E P V D V D C F C R N V D G V Y
601 TTGGTGTGAT GATAGCCTTT CTATGACTG CTAACCATTA GATCAAGGTG AGAACCTGT TGACCTTGAT TGCTTCTGCC GAAACGTGSA TGGGGTGTAT
   AACCCACTA CTATCGGAAA GARTACTCAC GCATTGGTAT CTAGTTCAC CCCTTGGACA ACTGCAACTA ACGAAGAGGG CTTTGCACCT ACCCCACATA
   prM protein
-----
   L E Y G R C G K Q E G S R T R R S V L I P S H A Q G E L T G R G H K
701 CTCGAATATG GACGGTGTG TAAACAAGAA GGAAGCAGAA CCAGACGCTC AGTGCCTATA CCTTCCACG CTCGAAGAGA GCTGACGGGA CGGGGACATA
   GAGCTTATAC CTGCCACACC ATTTGTCTT CTTTCTGCTT GGTCTGCGAG TCACGAATAT GGGAGGGTGC GACTTCTCTT CACTGGCCCT CCCCCTGTAT

```

prM protein

801 · W L E G D S L R T H L T R V E G W V W K N R L L A L A M V T V V W
AATGGTTGA GGGCGACTA CTCGGAACAC ATTTGACCCG CGTCGAGGGC TGGTCTGGA AAAATCGGCT GTTGGCCCTC GCTATGGTGA CAGTCGTTTG
TTACCAACCT CCCGCTGAGT GAGGCTTGTG TAAACTGGCG GCAGCTCCCG ACCCAGACCT TTTTAGCCGA CAACCGGGAG CGATACCACT GTCAGCAAAC
Hypr E protein

prM protein

901 · L T L E S V V T R V A V L V V L L C L A P V Y A S R C T H L E N R
GCTCAGCCTG GAGTCTGTGG TTA CTGGCTG GGCAGTGGCT GTGGTCTTC TGTCTTTG CCGTCTCTAC CGTCCCGGT GTACTCATTT GGAACAACAG
CGAGTGGAC CTCAGAGACC AATGAGCCCA CGTCAAGCA CCCCAGGAGG AGACAGAAGC GGGACAGATG CCGAGGTCCA CATGAGTAA CCTTTTGTCT
Hypr E protein

1001 · D F V T G T Q G T T R V T L V L E L G G C V T I T A E G K P S M D V
CTAAACACT GSCCTGGGT CCCCTGCTGA GCCCATTTGG ACCCAACT TGACCCACCA ACGCAATGAT AATGGCGACT CCCGTTTGGG AGATACCTAC
Hypr E protein

1101 · W L D A I Y Q E N P A Q T R E Y C L H A K L S D T K V A A R C P T
TGTCGCTGA TSCAATCTAT CAGGAAATC CCGCACAAAC CAGGGAATAT TGCCTTCAGC CAAAGCTGTC CGATACAAAG GTCGCGGCTA GGTGCCCAAC
ACACCCACT ACGTATGATA GTCCTCTAG GCGTGTGTTG GTCCCTTATA ACGGAAGTGC GTTTCGACAG CCTATGTTC CAGGCGCAT CCACGGTGTG
Hypr E protein

1201 · M G P A T L A E E H Q G G T V C K R D Q S D R G W G N H C G L F G
AATGGGACCG GCCACCCCTGG CCGAGGAACA TCAGGGAGGT ACAGTGTGCA AACGGGACCA GAGTGATAGA GGCTGGGTA ATCACTGGGG CCTGTTCGGC
TTACCTGGC CGTGGGACC GCTCCTTGT AGTCCCTCCA TGTCAACGCT TTGGCCCTGG CTCACTATCT CCGACCCCAT TAETGACGCC GGACAAGCCC
Hypr E protein

1301 · K G S I V A C V K A A C E A K R K A T G H V Y D A N K I V Y T V K V
AAAGGAAGTA TTGTCGCTTG CGTCAAGCA GCCTGTGAGG CCAAAAAGAA GGCTACTGGG CACCTCTATG ACGCCAACA GATCGTTTAT ACAGTAAAG
TTCTCTCAT AACAGCGAAC GCAGTTCGCT CGGACACTCC GGTTTTCTT CCGATGACCC GTGCAGATAC TCGGTTGTT CTAGCAATA TGTCACITTC
Hypr E protein

1401 · E P H T G C Y V A A N E T H S G R K T A S F T V S S E K T I L T M
TGGAACCACA CACAGGGAT TACGTGGCG CCAACGAGAC TCATTCCGGT CGAAAAACGG CCAGCTTAC CCGTCTATCC GAAAAGACCA TCCTCACTAT
ACCTGGTGT GTGTCCTTA ATGCACCGCC GTTTGTCTG AGTAAGGCCA GCGTTTGGCC GGTGGAAGTG GCACAGTAGG CTTTTCTGGT AGGAGTATA
Hypr E protein

1501 · G E Y G D V S L L C R V A S G V D L A Q T V I L E L D K T V E H L
GGGGGAGTAT GCGACGTTT CTCTGCTCTG CCGGGTGGCT AGCGGAGTGC ACCTGGCCCA GACAGTATC CTGGAAGTGG ATAAAACAGT TGAGCATCTG
CCCCCTATA CCGTCAAA GAGACGAGAC GGCCACCGA TCGCTCAGC TGGACCGGT CTGTAGTAG GACCTTGACC TATTTGTCA ACTCGTAGAC
Hypr E protein

1601 · P T A W Q V H R D W F N D L A L P W X H E G A R N W N N A E R L V E
CCTACCGCTT GCGAGTSCA CAGGATTTG TTTAAGACC TTGCCCTGCC ATGGAAACAT GAAGGAGCCA GAACTGGAA TAATGCAGAG CCACTCTAG
GGATGGCGAA CCGTCCAGT GTCCTAAC AAATTGCTG ACCGGACGG TACCTTTGTA CTCTCTGCT CTTTGACCTT ATTACGTCTC GCTGAGCATC
Hypr E protein

1701 · F G A P H A V K M D V Y N L G D Q T G V L L K A L A G V P V A H I
AATTCGGTGC CCTCATGCC GTGAGATGG ACGTCTACA TCTGGGTGAT CAGACCGCG TTCTCTTAA AGCTCTGCT GCGTCCAG TTGCCACAT
TTAAGCCAGG GGGAGTACG CACTTAC TGCAGATCT AGACCCACTA GTCGGCCG AGAGGAATT TCGAGAGCA CCGCATGCTC ACGGGGTGA
Hypr E protein

```

1801  * E G T K Y H L K S G H V T C E V G L E K L K M K G L T Y T M C D K
      CGAAGGAACG AAGTACCACC TGAAGTCAGG CCATGTAAC TGGAGGTGG GCCTGGAGAA GTTGAAAATG AAAGGTCTTA CGTACACAAAT GTGTGACAAAG
      GCTTCCITGC TTCATGTTGG ACTTCAGTCC GGTACATTGA ACGCTCCACC CGGACCTCIT CAACITTTAC TTCCAGAAAT GCATGTGTTA CACAGTGTTC
      Hypr E protein
      -----
1901  * T K F T W K R A P T D S G H D T V V M E V T F S G T K P C R I P V R
      ACCAAGTTCA CATGGAAGAG GGGCCCCACA GATAGCGGCC ACGTACTGT GGTGATGGAG GTGACCTTTT CTGGAACAAA ACCCTCCAGA ATACCCGTGC
      TGGITCAAGT GTACCTTCTC CCGGGGGTGT CTATCGCCGG TCGTATGACA CCATACCTC CACTGGAAAA GACCTTGTTT TGGGACGTCT TATGGGCACG
      Hypr E protein
      -----
2001  * A V A H G S P D V N V A M L I T F N P T I E N N G G G F I E M Q L
      GGGCTGTAGC TCACGGATCT CCCGATGTC ATGTTGCTAT GCTGATTACA CCTAACCCTA CCATCGAGAA TAACGGTGGT GGTITTTATG AGATGCAGCT
      CCCGACATCG AGTGCCTAGA GGGCTACAGT TACAACGATA CGACTAATGT GGATTGGGAT GGTAGCTCTT AITGCCACCA CCAAAAATAAC TCTACGTCCA
      Hypr E protein
      -----
2101  * P P G D N I I Y V G E L S Y Q W F Q K G S S I G R V F Q K T K K G
      TCCGCCAGGC GATAACATCA TCTACGTGGG CGAACTCTCT TACCAGTGGT TTCAGAAAAG GAGTCAATT GGGCGGTCT TCCAAAAAC GAAGAAGGA
      AGGGGGTCCG CTATTGTAGT AGATGCACCC GCTTGAGAGA ATGGTCACCA AAGCTTTCC CTCAAGTTAA CCCCACCAGA AGTTTTTTTG CTCTCTCCCT
      Hypr E protein
      -----
2201  * I E R L T V I G E H A W D F G S A G G F L S S I G K A L H T V L G G
      ATCGAACGAT TGACGGTTAT CGGCGAGCAC GCATGGGAT TGGTTCGGC AGGGGGATC CTGTCTTCTA TTGGTAAAGC ACTGCATACC GTGCTGGGG
      TAGCTTGCTA ACTGCCAATA GCGCTCGTG CGTACCCTAA AACCAAGCCG TCCCCTAAG GACAGAAGAT AACCATCCG TGACCTATGG CACGACCCCC
      Hypr E protein
      -----
2301  * A F N S I F G G V G F L P K L L L G V A L A W L G L N M R N P T M
      GCSCATCAA TTCTATTTTC GGGGCGTGG GGTTCCTGCC TAACTCCTG CTGGGAGTAG CCCTGGCCTG GTTGGGACTG AATATGGGA ATCCGACGAT
      CCGSTAAGTT AAGATAAAG CCCCAGCAC CCAAGGACGG ATTTGAGGAC GACCCTCATC GGGACCGGAC CAACCCTGAC TTATACGCCCT TAGGCTGCTA
      Hypr E protein
      -----
                                         NS1 gene of YF17D
2401  * S M S F L L A G V L V L A M T L G V G A D Q G C A I N F G K R E L
      GTCCATGTC TTCTCTTGG CCGGCGTGT TGTACTGGCC ATGACACTGG GCGTGGCCG CGATCAAGGA TGGCCCATCA ACTTTGGCAA GAGAGAGCTC
      CAGGTACAGT AAGGAGAACC GGGCCACCA ACATGACCGG TACTGTGACC CGCAACCGCG GCTAGTTCTT ACGCGGTAGT TGAACCCGTT CTCTCTCGAG
  
```

CV-TBEV Hypr with YFV/WNV chimeric signal (p45)

```

                                         5' UTR
1      AGTAAATCCT GTGTGCTAAT TGAGGTGCAT TGGTCTGCAA ATCGAGTTCG TAGGCAATRA ACACATTTGG ATTAAATTTA ATCGTTCGTT GAGCGATTAG
      TCATTTAGCA CACACGATTA ACTCCACGTA ACCAGACGTT TAGCTCAACG ATCCGTTATT TGTGTAACCC TAATTAATTA TAGCAAGCAA CTCGCTAATC
      5' UTR
      -----
                                         C protein YF17D
101   * M S G R K A Q G K T L G V N M V R R G V R S L S N K I K
      CAGAGAACTG ACCAGAACAT GCTGGTCTGT AAAGCTCAGG GAAAAACCTT GGGCGTCAAT ATGGTACGAC GAGGAGITCG CTCCTTGTC AAAAAATAA
      GTCTCTTGAC TGTCTTSTA CAGACCAGCA TTTCGAGTCC CTTTTGGGA CCGGAGTTA TACCATGCTG CTCCTCAAGC GAGGAACAGT TTGTTTTATT
      C protein YF17D
      -----
      * Q K T K Q I G N R P G P S R G V Q G F I F F F L P N I L T G K K I
  
```

201 AACAAAAAAC AAAACAAATT GGAACAGAC CTGGACCTTC AAGAGGTGTT CAAGGATTTA TCTTTTCTT TTTGTTCAAC ATTTTGACTG GAAAAAAGAT
 TTGTTTTTTG TTTTGTTPAA CCTTTGTCTG GACCTGGAAG TTCTCCACAA GTTCCTAAAT AGAAAAAGAA AAACAAGTTG TAAAAC TGAC CTTTTTTCTA
 C protein YF17D

· T A H L K R L W K M L D P R Q G L A V L R K V K R V V A S L M R G
 301 CACAGCCAC CTAAGAGGTG TGTGAAAAT GCTGGACCCA AGACAAGGCT TGGCTGTCT AAGGAAAGTC AAGAGAGTGG TGGCCAGTTT GATGAGAGG
 GTGTCGGGTG GATTCTCCA ACACCTTTA CGACCTGGT TCTGTCCGA ACCGACAAGA TTCCTTTCAG TTCTCTCACC ACCGGTCAAA CTACTCTCCT
 C protein YF17D WNV partial signal

YF 17D partial signal Hypr prM protein

L S S R K R R S H D V L T V Q F L I L G M L A C V G A A T V R K E R
 401 TTGTCCTCAA GGAACGGCG TTCCCATGAT GTTCTGACTG TGCAATTCCT AATTTTGGGC ATGCTGGCTT GTCTCGGAGC AGCTACCCCTG CGAAAGAGAC
 AACAGGACTT CCTTTGCCGC AAGGGTACTA CAAGACTGAC ACGTTAAGGA TTAACCCCG TACGACCGAA CACAGCCTCG TCGATGGCAC GCTTTTCTTG
 Hypr prM protein

· D G S T V I R A E G K D A A T C V R V E N G T C V I L A T D M G S
 501 GCGACGAGG CACCGTGATA AGGGCTGAGG GTAAGGATGC GGCTACCGCAG GTGAGAGTAG AGAATGGCAC TTGCGTAATA CTCGGCACTG ATATGGGATC
 CGCTGCCTTC GTGGCACTAT TCCCGACTCC CATTCTCAGC CCGATGGCTC CACTCTCATC TCTTACCGTG AACGATATAT GAGCGCTGAC TATACCTTAG
 Hypr prM protein

· W C D D S L S Y E C V T I D Q G E E P V D V D C F C R N V D G V Y
 601 CTGGGTGAC GATAGCCCTCA GTTATGAATG CGTAACAATA GACCAGGGCG AAGAACCCTGT GGACGTTGAC TGTTCCTGTA GAAATGTGGA TGGCGTTTAT
 ACCACACTG CTATCGGAGT CAIACTTAC GCATTGTTAT CTGCTCCCGC TTCTTGGACA CCGCAACTG ACAAGACAT CTTTACACTT ACCCGAATA
 Hypr prM protein

L E Y G R C G K Q E G S R T R R S V L I P S H A Q G E L T G R G H K
 701 CTGGAGTACG GCCGCTGTGG AAAACAGGAG GGCTCACGAA CTCGAAGATC TGTGCTGATT CCAAGTCACG CGCAAGGAGA GTTGACCGGT AGAGGCCACA
 GACCTCATG CCGGACACCC TTTTGTCTC CCGAGTGGTT GAGCTTCTAG ACAGACTAA GGTTCAGTGC GCGTTCCTCT CACTGGCCA TCTCCGGTGT
 Hypr prM protein

· W L E G D S L R T H L T R V E G W V W K N R L L A L A M V T V V W
 801 ACTGGCTTA AGGGACTCA TTGAGGACCC ACCTGACTAG GGTGGAGGGT TGGGTTTGGG ACAAATCGTT GCTCCGCTC GCTATGTTCA CCGTCTGTG
 TCACCAACT TCCCTGAGT AACTCTGGG TGGACTGATC CCACCTCCCA ACCCAACCT TCTTAGCCAA CGAGCGCGAG CGATACCAGT GGCAGCACAC
 Hypr prM protein

Hypr E protein

· L T L E S V V T R V A V L V V L L C L A P V Y A S R C T H L E N R
 901 GCTGACTG GAGAGTGTG TACTCGGGT TGCTGTGTTG GTTGTCTCC TCTGTTTGGC CCCAGTGTAC GCGTCCAGT GACTACTATT GGAACAACA
 CGACTGTGAC CTCTCACAG ACTGAGCCA ACACACAAC CAACAGGAGG AGACAAACCG GGTTCACATG CGCAGGTCCA CATGAGTAAA CCTTTTGTCT
 Hypr E protein

D F V T G T O G T T R V T L V L E L G G C V T I T A E G K P S M D V
 1001 GATTTTGTCA CCGGACCCA GGGGACGACT CGGGTAACCC TGSTGCTTGA ACTGGGTGGT TCCGTTACTA TTACCGCTGA GGGCAACCC TCTATGGATG
 CTAAACAGT GGCCGTGGT CCCCCTGCTA GCCCATGGG ACCAGCACTI TGACCCACCA ACCCAATGAT AATGGGACT CCGGTTTGGG AGATACCTAC
 Hypr E protein

· W L D A I Y Q E N F A Q T R E Y C L H A K L S D T K V A A R C P T
 1101 TGTGGTGGG TGCAATCTAT CAGGAGAATC CCGCACAAAC CAGGGAATAT TGCTTCACG CAAGCTGTG CGTACAAG GTCCGGCTA GGTGCCACAC
 ACACCGACT ACCTTAGATA GTCCCTCTAG GCGGTGTTG GTCCCTATA ASGGAAGTGC GTTTCGACAG GCTATGTTTC CAGCGCCGAT CCACGGGTTG
 Hypr E protein

· M G P A T L A E E H Q G G T V C K R D Q S D R G W G N H C G L F G

1201 AATGGGACCG GCCACCTGG CCGAGGAACA TCAGGGAGGT ACAGTGTGCA AACGGGACCA GAGTGATAGA GGCTGGGGTA ATCACTGCGG CCTGTTCGGC
 T'PACCTTGC CCGTGGGACC GCCTCCTTGT AGTCCCTCCA TGTACACAGT TTGCCCTGGT CTCACATATCT CCGACCCCAT TAGTGACGGC GGACAAGCCG
 Hypr E protein

K G S I V A C V K A A C E A K K K A T G H V Y D A N K I V Y T V K V
 1301 AAAGGAAGTA TTGTCGCTTG CGTCAAGGCA GCCTGTGAGG CCAAAAAGAA GGCTACTGGG CACGTCTATG ACCCAACAA GATCGTTTAT ACAGTGAAG
 TTTCTTCAT AACAGCGAAC GCAGTTCCTT CCGACACTCC GGTTTTCTT CCGATGACCC GTGCAGATAC TCGGGTTGTT CTAGCAAAATA TGTCACTTTC
 Hypr E protein

E P H T G D Y V A A N E T H S G R K T A S F T V S S E K T I L T M
 1401 TGGAACACA CACAGGGAT TACGTGGCGG CCAACGAGAC TCATTCGGGT CGCAAAACGG CCAGCTTAC CGTGTCTCC GAAAAGACCA TCCTCACTAT
 ACCTTGGTGT GTGTCCCTCA ATGCACCGCC GGTTCCTCIG AGTAAGGCCA GGTTTTGGC GGTGCAAGTG GCACAGTAGG CTTTCTGGT AGGAGTGATA
 Hypr E protein

G E Y G D V S L L C R V A S G V D L A Q T V I L E L D K T V E H L
 1501 GGGGAGTAT GGCAGGTTT CTCTGCTCTG CCGGGTGGT ACGGGAGTGC ACCTGGCCCA GACAGTCAFC CTGCAACTGC ATAAACAGT TGAGCATCTG
 CCCCCTATA CCGTCCAAA GAGAGGAGC GGCCCAACGA TCCCTTCACC TGGACCGGT CTGTCAAGTG GACCTTGACC TATTTTGTCA ACTCGTAGAC
 Hypr E protein

P T A W Q V H R D W F N D L A L P W K H E G A R N W N N A E R L V E
 1601 CCTACCGTT GGCAGTGCA CAGGGATTGG TTTAACGACC TTGCCCTGCC ATGGAACAT GAAGGAGCGA GAAACTGGAA TAATGCAGAG CGACTCGTAG
 GGATGGCGAA CCGTCCAGT GTCCCTAACC AAATTGCTGG AACGGGAGCG TACCTTTGTA CTTCCTCGCT CTTTGACCT ATTGACCTC GCTGAGCATC
 Hypr E protein

F G A P H A V K M D V Y N L G D Q T G V L L K A L A G V P V A H I
 1701 AATTCGGTGC CCTCATGCC GTGAAGATGG ACGTCTACAA TCTGGGTGAT CAGACCGGGC TTCTCCTTAA AGCTCTCGCT GGCGTACCAG TTGCCACAT
 TFAAGCCAGC GGGAGTACGG CACTTCTACC TGCAGATGT AGACCCACTA GTCTGGCCGC AAGAGGAAT TCGAGAGCGA CCGCATGCTC AACGGGTGTA
 Hypr E protein

E G T K Y H L K S G H V T C E V G L E K L K M K G L T Y T M C D K
 1801 CGAAGAACG AATACCACC TGAAGTCAGG CCATGTAAT TCGAGGTGG GCCTGGAGAA GTTGAAAATC AAAGGTCTTA CGTACACAAT GTGTGCAAG
 GCTTCCTTGC TTCATGGTGG ACTTCAGTCC GGTACATTGA ACGCTCCACC CCGGACCTTT CAACCTTTTAC TTTCCAGAAI GCATGTGTTA CACACTGTTC
 Hypr E protein

T K F T W K R A P T D S G H D T V V M E V T F S G T K P C R I P V R
 1901 ACCAAGTCA CATGGAAGAG GGCCCCACA CATACCGGCC ACGATACTGT GGTGATGGAG GTGACCTTTT CTGGAACAAA ACCCTGCAGA ATACCCGTGC
 TGGTCAAAT GTACCTTCTC CCGGGGTGT CTATCGCCCG TGCTATGACA CCACTACCTC CACTGGAAAA GACCTTGTTT TGGGACCTCT TATGGGCAGC
 Hypr E protein

A V A H G S P D V N V A M L I T P N P T I E N N G G G F I E M Q L
 2001 GGGCTGTAGC TCACGGATCT CCGATGICA ATGTTGCTAT GCTATTACA CCTAACCTTA CCATCGAGAA TAACGGTGGT GGTTTTATTG AGATGCAGCT
 CCCGACATCG AGTGCCTAGA GGGCIACAGT TACAACGATA CCACTAATGT GGATGGGAT GGTAGCTCTT ATTGCCACCA CAAAATAAC TCTACCTCGA
 Hypr E protein

P P G D N I I Y V G E L S Y Q W F Q K G S S I G R V F Q K T K K G
 2101 TCCGCCAGG GATAACATCA TCTACGTTGG CCAACTCTCT TACAGTGGT TTCAGAAAGG GAGTCAATT GSGCGGGTCT TCCAAAAAC GAACAGCGGA
 AGCCGGTCCG CTATTGTAGT AGATGCAACC GCTTGAGAGA ATGCTCACCA AAGCTTTTCC CTCAGTTAA CCGGCCAGA AGGTTTTTTG CTTCCTCCCT
 Hypr E protein

I E R L T V I G E H A W D F G S A G G F L S S I G K A L H T V L G G
 2201 ATCGAACGAT TGACGGTAT CCGGAGCAC GCATGGGATT TTGGTTCGC AGGGGATTC CTGTCTTCTA TTGGTAAGGC ACTGCATACC GTGCTGGGG
 TAGCTTGCTA ACTGCCAATA GCCGTCTGTG CGTACCCCTAA AACCAAGGCG TCCCCCTAAG GACAGAAGAT AACCATTCG TGACGTATGG CACGACCCCC
 Hypr E protein

2301 · A F N S I F G G V G F L P K L L L G V A L A W L G L N M R N P T M ·
 GCGCAITCAA TTCIATITTC GGGGGCGTGG GGTTCCTGCC TAAACTCCIG CTGGGAGTAG CCCTGGCCTG GTTGGGACTG AATATGCGGA ATCCGACGAT
 CGCCTAAGTT AAGATAAAGC CCCCCGACC CCAAGGACGG ATTTGAGGAC GACCCTCATC GGGACCGGAC CAACCCCTGAC TTATACGCCT TAGGCTGCTA
 Hypr E protein

NS1 gene of YF17D
 2401 · S M S F L L A G V I V L A M T L G V G A D Q G C A I N F G K R E L ·
 GTCCATGTCA TTCCTCTTGG CCGGCGTGCT TGTACTGGCC ATACACTGG GCGTTGGCGC CGATCAAGGA TGCGCCATCA ACTTTGGCAA GAGAGAGCTC
 CAGGTACAGT AAGGAGAACC GGCCCGACGA ACATGACCGG TACTGTGACC CGCAACCGCG GCTAGTTCCT ACGCGGTAGT TGAACCGGT CTCTCTCGAG

CV-LGTV E5 with YFV/TBEV chimeric signal (p43)

5' UTR
 1 AGTAAATCCT GTGTGCTAAT TGAGGTGCAT TGCTCTGCAA ATCGAGTTGC T'AGGCAATAA ACACATTGG ATTAATTTTA ATCCTTCGTT GAGCGATTAG
 TCATTAGGA CACACAGTA ACITCACGTA ACCAGACGTT TAGCTCAACG ATCCGTTATT TCTGTAAACC TAATTAATAAT TAGCAAGCAA CTCGCTAATC
 5' UTR

C protein YF17D
 101 M S G R K A Q G K T L G V N M V R R G V R S L S N K I K ·
 CAGAGAAGCTG ACCAGAACAT GTCTGGTCTGT AAGCTCAGG GAATAACCCG GGGCTCAAT ATGCTACGAC GAGGAGTTCG CTCCTTGTCA AACAAATAA
 GTCTCTTGAC TGGTCTTSTA CAGACCACCA TTTCGAGTCC CTTTTTGGGA CCGCGAGTTA TACCATGCTG CTCCTCAAGC GAGGAACAGT TTGTTTTATT
 C protein YF17D

201 · Q K T K Q I G N R P G P S R G V Q G F I F F F L F N I L T G K K I ·
 AACAAAAAC AAAACAATTT GGAACAGAC CTGGACCTTC AAGAGGTGTT CAAGGATTTA TCTTTTCTT TTTGTTCAAC ATTTTGACTG GAATAAAGAT
 TTGTTTTTGT TTTTGTITIA CCTTTGTCTG GACCTGGAAG TTCTCCACAA GTTCCCTAAT AGAAAAAGAA AAACAAGTTC TAAACTGAC CTTTTTCTA
 C protein YF17D

301 · T A H L K R L W K M L D P R Q G L A V L R K V K R V V A S L M R G ·
 CACAGCCAC CTAAGAGGTG TGTGCATAAT GCTGCACCCA AGCAGGGCT TGGGTGTTCT AAGGAAGTC AAGAGACTGG TGGCCAGTTT GATGAGGGA
 GTGTCGGGTG GATTTCTCCA ACACCTTTTA CGACCTGGGT TCTGTTCCGA ACCGACAAGA TTCCPTTCAG TTCTCTCACC ACCCGTCAA CTACTCTCCT
 C protein YF17D TBEV partial signal

YF 17D partial signal prM protein Langat E5
 401 L S S R K R R S H D V L T V Q F L I L G M L G M T I A A T V R R E R ·
 TTGCTCTCAA GCAACGCGCC TTCCCATGAT GTCTGACTG TGCAATTCCT AATTTTGGGC ATGCTGGGGA TGACGATCGC AGCTACTGTG CGAAGGGAGA
 AACAGGAGTT CCTTTGCGGC AAGGGTACTA CAAGACTGAC ACCTTAAGGA TTAACCCCG TACGACCCCT ACTGCTAGCG TCGATGACAC GCITCCCTCT
 prM protein Langat E5

501 · D G S M V I R A E G R D A A T Q V R V E N G T C V I L A T D M G S ·
 GAGACGCTC TATGGTATC AGAGCCGAAG GTAGGGACGC TGCAGCCGAG CTGAGGTCG AAATGGCAC CTCTGTATT CTGCGACCG ACATGGGCTC
 CTCTGCCGAG ATACCACTAG TCTCGGCTTC CATCCCTGCG ACGCTGGCT CACTCCGAG TTTTACCCTG GACACATAA GAGGCTGCG TGTACCCGAG
 prM protein Langat E5

601 · W C D D S L A Y E C V T I D O G E E P V D V D C F C R G V E K V T ·
 CTGGTGTGAT GATTCTCTGG CTTATGAATG TGTACTATT GATCAGGGTG AAGAGCCTGT GCACGTGGAC TGTTCCTGTA GAGGCTCGA GAAAGTGACC
 GACCACACTA CTAAGAGACC GAATACTTAC ACAATGATAA CTAGTCCCAC TTCTCGGACA CCTGCACCTG ACAAAGACAT CTCCGAGCT CTTTCACTGG

prM protein Langat E5

701 L E Y G R C G R R E G S R S R R S V L I P S H A Q R D L T G R G H Q
 CTGGAATATG GACGATGTGG CCGGCGAGAA GGCTCCAGGA GTCCGGAGATC CGTGTGTATC CCTTCACATG CGCAGCGCGA TCTGACAGGG AGGGGTCCAC
 GACCTTATAC CTGCTACACC GGGCGCTCTT CCGAGGTCTT CAGCCTCTAG GCACAACTAG GGAAGTSTAC GCGTCGCGCT AGACTGTCCC TCCCCAGTGG

prM protein Langat E5

801 W L E G E A V K A H L T R V E G W V W K N K L F T L S L V M V A W
 AGTGGCTCGA AGGCGAAGCA GTCRAAGGCC ATCTGACTCG CGTTGAAGGC TGGGTGTGGA AAACAAACT CTTTACCTT AGCCTGGTGA TGGTCGGTG
 TCACCAGCT TCCCTCCGT CAGTCCGGG TAGACTGAGG GCACCTCCG ACCCACCT TTTTGTGTA GAAATGGAA TCGGACCACT ACCAGCCAC
 prM protein Langat E5

E protein Langat E5

901 L M V D G L L F R I L I V V V A L A L A P A Y A S R C T H L E N R
 GCTGATGTA GACGACTCC TTCGCCCAT TCTCATTTGT GTGGTGGTC TCGCGCTCG CCTTCACATC GGTCCAGGT GTACGCACCT CGAAAATCGA
 CGACTACCAT CTGCTGAGG AAGGGCGTA AGAGTAACA CACCACCAG AGCGAGAGG GGGACGTATG CGCAGTCCA CATGCGTGA GCTTTAGCT
 E protein Langat E5

1001 D F V T G V Q G T T R L T L V L E L G G C V T V T A D G K P S L D V
 GATTTCGCA CAGCGTCCA AGTACTACC CGCTCACCC TCGTGTGGA GGTGGAGGC TGTGTCACTG TTACAGCCGA CGGAAAACCT AGTCTGGATG
 CTAAGCAGT GTCCGAGGT TCCATGATGG GCCGAGTGG AGCAGGACT CGACCCCTCG ACACAGTAC AATGTCCGCT GCCTTTTGA TCAGACCTAC
 E protein Langat E5

1101 W L D S I Y Q E S P A Q T R E Y C L H A K L T G T K V A A R C P T
 TGTGGCTGA CTCCATCTAT CAGGAGAGCC CGGCACAGC CAGGGACTAC TGCCCTCACG CTAAGCTGAC TGGACAAAG GTAGCCGAA GATGTCCAC
 ACACCGACT GAGGTAGATA GTCTCTCG GCCGTGTCT GTCCCTCATG ACGGAGTGG GATTGCACTG ACCCTGTTC CATCGGCTT CTACAGGTTG
 E protein Langat E5

1201 M G P A T L P E E H Q S G T V C K R D Q S D R G W G N H C G L F G
 AATGGGCCCT GCCACCTTG CCGAGGAACA CCAATCCGGT ACGGTATGCA AGCGAGATCA GTCTGATCG GGTGGGGGA ATCATTCGG CCTCTCGGT
 TTACCCCGGA CGGTGGAAGG GGCTCCTGT GGTAGGCCA TGCCATAGT TGGTCTAGT CAGACTAGCG CTRCCCTCT TAGTAACGCC GGAGAGCCA
 E protein Langat E5

1301 K G S I V T C V K V T C E D K K K A T G H V Y D V N K I T Y T I K V
 AAAGGCACA TTGTCCTTG CTTGAGGTG ACATCCGAGG ACAAGAAGA GGCCACAGGT CATGTATATG ATGTGAACA AATCAGATAT ACCATTAAG
 TTTCCGCTG AACAGTGAAC GCACCTCCAC TGTACGCTCC TGTCTTCTT CCGGTGTCCA GTACATATAC TACACTTGT TTAGTGTATA TGGTAATTCC
 E protein Langat E5

1401 E P H T G E F V A A N E T H S G R K S A S F T V S S E K T I L T L
 TAGARCCCA TACAGGGAA TTCGTGGCAG CAAACGAGAC TCATAGCGGA CGAAAGTCCG CCTCCTTAC CGTCTCCTCC GAGAAAACA TCCTGACCT
 ATCTTGGTGT ATGCCCCTT AAGCACCTC GTTGTCTCTG AGTATCGCT GCTTTCAGC GGAGGAAGTG GCACAGGAGG CTCTTTTGT AGGACTGGGA
 E protein Langat E5

1501 G D Y G D V S L L C R V A S G V D L A Q T V V L A L D K T H E H L
 CGGAGACTAC GCGACATAT CTTTGTGTG CAGGGTGGC AGCGGCTGG ACCTTGTCA GACAGTCTG TTGGCCCTGG ACAAGACACA TGAGCACTT
 GCCTCTGATG CCGCTGATA GAAACGACAC GTCCACCGG TCGCCGACC TGGACGAGT CTCTCAGCAC AACCGSACC TCTTCTGT ACTCGTGAAC
 E protein Langat E5

1601 P T A W Q V E R D W F N D L A L P W K H D G A E A W N E A G R L V E
 CCAACAGCT GGCAGGTGA CAGGACTGG TTTRACGACC TGGCGTCCC GTGGAAACAT GACGGCGCTG AAGCATGAA TGAGGCAGG AGACTGTGG
 GGTTCGGA CCGTCCACT GTCCCTGACC AAATGCTGG ACCGCGAGG CACTTTGTA CTCGCCGAC TTCCCTTCTT ACTCCGTC TCTGACCAC
 E protein Langat E5

```

· F G T F H A V K M D V F N L G D O T G V L L K S L A G V P V A S I ·
1701 AATTTGGAC CCCACAGCC GTAAAGATGG ACGTITTCAA TCTTGGTGAC CAGACAGGGG TGCTCCTGAA ATCACTGGCG GGCGTGCCCTG TAGCCAGCAT
TTAARCCCTG GGGTGTGCGG CATTTCCTACC TGCAAAAAGTT AGAACCACCTG GTCCTGCCCC ACGAGGACTT TAGTGACCCG CGGCACGGAC ATCGGGCTGA
E protein Langat E5
-----
· E G T K Y H L K S G H V T C E V G L E K L K M K G L T Y T V C D K
1801 CGAGGGCACA AAGTATCAC TGAAGTCTGG GCATGTAACC TGGGAAATGG GCCTGGAAAA GCTCAAGATG AAAGGACTTA CGTACACTGT TTGIGATAAG
GCTCCCCGTG TTCATAGTGG ACTTCAGACC CSTACATTGG ACGCTPCACC CGGACCTTTT CGACTTCTAC TTTCCTGAAT GCATGTGACA AACACTATTG
E protein Langat E5
-----
T K F T W K R A P T D S G H D T V V M E V G F S G T R P C R I P V R ·
1901 ACCAAGITTA CATGGAAGCG AGCCCCAACC GATTCCGGCC ATGATACCGT CGTGATGGAG GTTGGTTTCT CCGGCACCAG ACCATGTAGA ATACCAGTGA
TGGTTCAAAT GTACCTTGGC TCGGGGTTGC CTAAGGCCGG TACTATGGCA GCACCTACCTC CAACCAAGA GCGCGTGGTC TGGTACATCT PATGGTCACT
E protein Langat E5
-----
· A V A H G V P E V N V A M L I T P N P T M E N N N G G G F I E M Q L ·
2001 SAGCTGTGCG CCACGCTGTA CCCGAGGTAA ACGTGGCCAT GCTGATTACA CCGAATCCCA CTATGGAGAA CAATGGCCGA GGCTTCATCG AAATGCAGCT
CTCGACAGCG GGTGCCACAT GGGCTCCATT TGCACCGGTA CGACTAATGT CGCTTAGGGT GATACCTCTT GTTACCGCCT CCCAAGTAGC TTTACCTCGA
E protein Langat E5
-----
· P P G D N I I Y V G D L D H Q W F Q K G S S I G R V L Q K T R X G
2101 GCCGCTCGA GACAACATCA TTTATGTCGG CGACCTCGAT CATCAATGGT TCCAGAAAGG GTCTCCATC GGCCCGCTCC TTCAGAAGAC ACCAAAAGGC
CGCCGGACCT CTGTTTGTAG AAATACAGCC GCTGGAGCTA GTAGTTACCA AGGTCTTTCC CAGAAGGTAG CCGCCGACAG AAGTCTTCTG TGCITTTCCG
E protein Langat E5
-----
I E R L T V L G E H A W D F G S V G G V M T S I G R A M H T V L G G ·
2201 ATTGAAAGAC TTACAGTCT GGGCGAACAT GCCTGGGACT TCGGGTCACT TGGCGGGTGA ATGACAAGGA TAGGCAGACC TATGCACACC GTTCTCGGTG
TAACTTTCTG AATGTCAGGA CCCGCTTGTA CGGACCCCTG AGCCCACTCA ACCGCCCAT TACTGTTCGT ATCCGTCTCG ATACGTGTGG CAAGAGCCAC
E protein Langat E5
-----
· A E N T L L G G V G F L P K I L L G V A M A W L G L N M R N P T L ·
2301 GGGCATTAA TACICIGTTG GGTGGCGTGG GTTTCTTCC GAAAATCCTG CTCGGTGTG CAATGGCCIG GCTTGGACTG AATATGGCCA ATCCTACACT
CCCCTAAAT ATGAGACAAC CCACCGCAC CAAAAGAAGG CTTTATAGAC GAGCCACAGC GTTACCGGAC CGAACCTGAC TTATACCGCT TAGGATGTGA
E protein Langat E5
-----
NS1 gene of YF17D
-----
· S M G F L L S G G L V L A M T L G V G A D Q G C A I N F G K R E L
2401 GAGTATGGG TTCTCTCTGT CAGGAGCCCT GGTCTTGGA ATGACTCTGG GAGTGGGCGC CGATCAAGGA TGGCCATCA ACTTTGGCAA GAGAGACCTC
CTCATACCCC AAAGAAGACA GTCCTCCGGA CCAGGACCTT TACTGAGACC CTCACCCCGG GCTAGTTCCT ACGCGGTAGT TGAACCGGT CTCTCTCGAG

```

CV-TBEV Hypr with YFV/TBEV chimeric signal and dC2 deletion in C protein (p59)

```

5' UTR
-----
1 AGTAAATCC1 GTGTGCTAAT TGAGGTGCAT TCGTCTGCAA ATCGAGTTCG TAGGCAATAA ACACATTGG ATTAATTTTA ATCGTTCGTT GAGCGATTAG
TCATTTAGGA CACACGATTA ACTCCACGTA ACCAGAGCCT TAGCTCAACG ATCCGTTATT TGTGTAACC TAATTAATAA TAGCAAGCAA CTCGCTAATC
5' UTR
-----
C protein
-----

```

101 CAGAGAAGCTG ACCACAACAT GTCTGGTGGT AAAGCTCAGG GAAAAACCCCT GGGCGTCAAT ATGGTACGAC GAGGAGTTCG CTCCITGTC AAAAAATAA
 GTCCTTTGAC TGGTCTTGTG CAGACCAGCA TTTCGAGTCC CTTTTGGGA CCGCAGTTA TACCATGCTC CTCCTCAAGC GAGGAACAGT TTGTTTTATT
 dC2 deletion (PSR)

C protein

201 Q K T K Q I G N R P G G V Q G F I F F F L F N I L T G K K I T A H
 AACAAAAAC AAAACAATT GGAACAGAC CTGGAGGTGT TCAAGGATT ATCTTTTCT TTTTGTCAA CATTTCAGT GGAABAAGA TCACAGCCCA
 TGTCTTTTGG TTTTGTAAA CCTTTGCTG GACCTCCACA AGTTCCTAAA TAGAAAAAGA AAAACAAGT GTAAAAGTA CCTTTTTTCT AGTGTCCGGT

C protein

301 L K R L W K M L D P R Q G L A V L R K V K R V V A S L M R G L S S
 CCTAAAGAG TTGTGAAAA TGCTGGACCC AAGACAAGG TTGGCTGTC TAAGGAAAGT CAAGAGAGT GTGGCCAGT TGATGAGAG ATTGTCTCCA
 GGATTTCTCC AACACCTTT ACAGCTGGG TTCTGTCCG AACCGACAAG ATTCCTTCA GTTCTCTCAC CACCGGTCAA ACTACTCTCC TAACAGGAGT

YF17D partial signal

TBEV partial signal

C protein

Hypr prM protein

401 R K R R S H D V L T V Q F L I L G M L G M T I A A T V R K E R D G S
 AGGAAGCCG GTTCCCATGA TGTCTGACT GTGCAATICC TAATTTTGGG CATGCTGGGC ATGCAATCG CAGTACCGT TCGCAAGGAA AGAGACGGCA
 TCCTTTCCGG CAAGGCTACT ACAAGACTGA CACGTTAAGG ATTAACAACC GTACGACCCG TACTGTTAGC GTGATGCCA AGCGTTCCCT TCTTGCCTG

Hypr prM protein

501 T V I R A E G K D A A T Q V R V E N G T C V I L A T D M G S W C D
 GTACGGTCAT ACGCGGGAA GGTAAAGGAG CCGCTACCCA AGTGAGAGTG GAAAAAGTA CCTGCCTCAT TCTGCCACC GACATGGCT CITGGTGTGA
 CATGCCAGTA TCGCCGCCCT CCATTCCTAC GCGGATGGGT TCACITCTCAC CTTTTACCAT GGACGCAGTA AGACCGGTGG CTGTACCCGA GAACCAACT

Hypr prM protein

601 D S L S Y E C V T I D Q G E E P V D V D C F C R N V D G V Y L E Y
 TGATAGCCTT TCTTATGAGT GCGTAACCAT AGATCAAGGT GAGGAACCTG TTGACGTTGA TTGCTTCTGC CGAAACGTGG ATGGGTGTA TCTCGMATAT
 ACTATCGGAA AGAATACTCA CGCATCGTA ICTAGTCCA CTCTTGGAC AACTGCAACT AACGAAGAG GCTTTCACCC TACCCACAT AGAGCTTATA

Hypr prM protein

701 G R C G K Q E G S R T R R S V L I P S H A Q G E L T G R G H K W L E
 GGACGCTGTG GTAACAAGA AGGAAGCAGA ACCAGACCGT CAGTCTTAT ACCCTCCAC GGTCAAGSAG AGCTGACCGG ACGGGACAT AAATGGTTGG
 CCGCCACAC CMTTGTCT TCTTCTCT TGGTCTGCGA GTCAAGGATA TGGAGGGTG CGAGTCTCT TCGACTGGCC TCGCCCTGTA TTTACCAACC

Hypr prM protein

801 G D S L R T H L T R V E G W V W K N R L L A L A M V T V V W L T L
 AGGGCGACTC ACTCCGAACA CATTGACCC GCGTCGAGGG CTGGGTCTGG AAAAAACGGC TGTGGCCCT CGCTATGGTG ACAGTCGTTT GGCTCACGCT
 TCCCCTGAG TGAGGCITGT GTAACCTGGG CGCAGCTCCC GACCCAGACC TTTTACCGG ACAACCGGGA GCGATACCAC TGTCAGCAAA CCGAGTCCGA

Hypr E protein

Hypr prM protein

901 E S V V T R V A V L V V L L C L A P V Y A S R C T H L E N R D F V
 GGAGTCTGTG STTACTGCGS TGGCAGTGT GGTGGTGTCT CTCTGCTTG CCCGTGCTA CGCGTCCAGG TGTACTCAT TGGAAAAACAG AGATTTTGT
 CCTCACACAC CAATGAGCGC ACCGTACGA CCACCCAGAG GAGACAGAAC GGGACAGAT GCGCAGTCC ACATGAGTAA ACCTTTTGTCT TCTAAAACAG

Hypr E protein

T G T Q G T T R V T L V L E L G G C V T I T A E G K P S M D V W L D

1001 ACCGGCACC AGGGGACGAC TCGGGTAACC CTGGTGCTTG AACTGGGTGG TTGGCTTACT ATTACCGCTG AGGGCAAACC CTCTATGGAT GTGTGGCTGG
 TGGCCGTGGG TCCCTCTGCTG AGCCCATTTGG GACCACGAACT TTAGCCACC ACCCAATGA TAATGGCGAC TCCCGTTTGG GAGATACCTA CACACCCGACC
 Hypr E protein

· A I Y Q E N P A Q T R E Y C L H A K L S D T K V A A R C P T M G P ·
 1101 ATCCAATCTA TCAGGAGAAT CCGGCACAAA CCAGGGGATA TTGCCTTAC GCAAAGCTGT CCGATACAAA GGTGCGGGT AGTGCCCAA CAATGGGACC
 TAGCTTAGAT AGTCTCTTA GGGCGTGTG GTTCCCTTAT AACGGAAAGT CTTCCGACA GCCTATGTTT CCGAGCGCGA TCCACGGGTT GTTACCCCTGG
 Hypr E protein

· A T L A E E H Q G G T V C K R D Q S D R G W G N H C G L F G K G S ·
 1201 GCGCCCTGT GCGGAGAAC ATCAGGCGAG TACAGTGTGC AAACGGGACC AGAGTGATAG AGGCTGGGGT AATCACTGGG GCCTGTTCGG CAAGGGAAGT
 CCGGTGGGAC GCGCTCCCTG TAGTCCCTCC ATGTACACAG TTTCCTCTGG TCTCACTATC TCCGACCCCA TTAGTGACCG CCGACAAGCC GTTCCCTTCA
 Hypr E protein

· I V A C V K A A C E A K K K A T G H V Y D A N K I V Y T V K V E P H ·
 1301 ATTGTCCCTT GCGTCAAGGC AGCCTGTGAG GCCAAAAGA AGGCTACTGG GCACGTCTAT GACGCCAACA AGATCGTTTA TACAGTAAA GTGGAAACCAC
 TAACAGCGAA CCGAGTTCG TCGGACACTC CCGTTTTTCT TCCGATGACC CGTGCAGATA CTGCGGTGTG TCTAGCAAT ATGTACTTTT CACCTTGGTG
 Hypr E protein

· T G D Y V A A N E T H S G R K T A S F T V S S E K T I L T M G E Y ·
 1401 ACACAGGGA TTAGTGGGG GCCAACGAGA CTCATTCCGG TCGCAAAAGC CCGAGCTCA CCGTGTATC CGAAAAGACC ATCCTCACTA TGGGGGASTA
 TGTGTCCCTT AATGCACCCG CCGTGTCTCT GAGTAAAGCC AGCCTTTTCC CCGTCGAAGT GGCACAGTAG GCTTTTCTGG TAGGAGTAT ACCCCCTCAT
 Hypr E protein

· S D V S L L C R V A S G V D L A Q T V I L E L D K T V E H L P T A ·
 1501 TGGCGACGTT TCTCTGCTCT GCGGGGTGGC TAGCGGAGTC GACCTGGCC AGACAGTCA TCTGGAAC TGATAAAACAG TTGAGCATCT GCCTACCGCT
 ACCGCTGCAA AGAGACGAGA CCGCCACCCG ATCCGCTCAG CTGGACCGGG TCTGTCACTA GGACCTTGAC CTATTTTGTG AACTCGTASA CCGATGGCGA
 Hypr E protein

· W Q V H R D W F N D L A L P W K H E G A R N W N N A E R L V E F G A ·
 1601 TGGCAGGTGC ACAGGGATTG GTTTAACGAC CTTGCCCTGC CATGGAACA TGAAGGAGCG AGAACTGGA ATATGCAGA GCGACTCSTA GAATTCGGTG
 ACCGTCCACG TGTCCCTAAC CAATTCCTG GAACGGGACG GTACCTTTGT ACTTCTCTCG TCTTTGACCT TATTACGCTT CCGTGAGCAI CTTAAGCCAC
 Hypr E protein

· F H A V K M D V Y N L G D Q T G V L L K A L A S V P V A H I E G T ·
 1701 CCCCCTATGC CGTGAAGATG GACGTCTACA ATCTGGGTGA TCAGACCGGC GTTCTCCTTA AAGCTCTCGC TGGCGTACCA GTTGCCCA CA TCGAAGGAAC
 GGGGAGIACG GCACCTTCTC CTGCAGATGT TAGACCCACT AGCTGTGGCC CAACAGGAAT TTGAGAGCG ACCGCATGTT CAACGGGTGT AGCTTCTTGG
 Hypr E protein

· K Y H L K S G H V T C E V G L E K L K M K G L T Y T M C D K T K F ·
 1801 GAAGTACCAC CTGAAGTCAG GCCATGTAAC TTGCGAGGTG GGCCTGGAGA AGTTGAAAAT GAAAGGTCTT ACGTACACAA TGTGTGACAA GACCAAGTTC
 CTTTCATGGT GACTTCAGTC CCGTACATTG AACGCTCCAC CCGGACCTT TCAACTTTTA CTTTCCAGAA TGCATGTGTT ACACACTGTT CTGGTTCAAG
 Hypr E protein

· T W K R A P T D S G H D T V V M E V T F S G T K P C R I P V R A V A ·
 1901 ACA7GGAAGA GGGCCCCAC AGATAGCGGC CACGATACIG TGGTATGGA GTGACCTTT TCTGGAACA AACCCCTGCG AATACCCGTG CCGCTCTAG
 TGTACCTTCT CCGGGGGGTG TCTATCGCCG CTGCTATGAC ACCACTAOCCT CCAC7GAAA AGACCTTGTT TTGGGACGTC TTATGGSCAC GCCCGACATC
 Hypr E protein

· H G S P D V N V A M L I T P N P T I E N N G S G F I E M Q L P P G ·
 2001 CTCACGATC TCCCGATGTC AATGTGCTA TGCTGATTAC ACCTAACCTT ACCATCGAGA ATAACGGTGS TGGTTTATT GAGATGCAGC TTCCGCCAGG
 GAGTGCCTAG AGGGCTACAG TTACAACGAT ACGACTAATG TGGATTGSSA TGGTAGCTCT TATTGCCAGC ACCAAAATAA CTCTACGTCG AAGCGCGTCC
 Hypr E protein

```

· D N I I Y V G E L S Y Q W F Q K G S S I G R V F Q K T K K G I E R
2101 CGATAACATC ATCTACGTTG GCGAACTCTC TTACCAGTGG TTTCAGAAAG GGAGTTC AAT TGGGCGGGTC TTCCAAAAA CGAAGAAGGG AATCGAACGA
GCTATTGTAG TAGATGCACC CGCTTGAGAG AATGGTCCACC AAAGTCTTTC CCTCAGTTA ACCCGCCAG AAGGTTTTTT GCTTCTCC CTTAGCTTGTCT
Hypr E protein
-----
L T V I G E H A W D F G S A G G F L S S I G K A L H T V L G G A F N ·
2201 TTGACGGTTA TCGGCGAGCA CGCATGGGAT TTGGTTCCG CAGGGGATTT CCTCTTCTCT ATTGGTAAGG CACTGCATAC CGTCTGGGG GGCGCATTC A
AACTGCCAAT AGCCGCTCGT CGTACCCCTA AAACCAAGCG GTCCCTTAA GGACAGAAGA TAACCATTCG GTGACGTATG GCACGACCCC CCGCGTAAGT
Hypr E protein
-----
· S I F G G V G F L P K L L L G V A L A W L G L N M R N P T M S M S ·
2301 ATTCTATTTT CGGGGCGGTG GGGTTCCTGC CTAAACTCCT GCTGGGAGTA GCCCTGGCTT GGTGGGACT GAATATGCGG AATCCGACGA TGTCATGTC
TAAGATAAAA GCCCCCGCAC CCCAAGGACG GATTTGAGGA CGACCTCAT CGGACCCGGA CCAACCTGA CTTATACGCC TTAGGCTGTG ACAGGTACAG
Hypr E protein
-----
NS1 gene of YF17D
-----
· F L L A G V L V L A M T L G V G A D Q G C A I N F G K R E L
2401 ATTCCTTCTG GCGGCGGTG TTGTACTGCG CATGACACTG GCGTGGGGC CCGATCAAGG ATGCGCCATC AACTTTGGCA AGAGAGAGCT C
TARGGAGAAC GCGCCGACG AACATGACC GTACTGTGAC CCGCAACCGG GGCTAGTTCC TACCGGTTAG TTGAACCGT TCTCTCTCGA G

```

Sequence Appendix 3

PIV-WN/TBEV Hypr with TBEV signal (p39)

```

deleted C
~~~~~
5' UTR
-----
1 AGTACTTCGC CTGTGTGAGC TGACAACTT AGTAGTCTTT GTGAGGATTA ACAACAATTA ACACAGTCCG AGCTGTTTCT TAGCACGAAG ATCTCGATGT
TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCCTAAT TGTTGTTAAT TGTGTACCGC TCGACAAAGA ATCTGTCTTC TAGAGCTACA
M S ·
WNV deleted C protein
-----
· K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K R S S K ·
101 CTAAGAAACC AGGAGGGCCC GSCAAGGCC GGGCTGTCTA TTTCGTAATA CGCGGAATGC CCCGCGTGT GTCCCTGATT GGACTTARGG GGAGCTCCAA
GATTCITTTG TCCPCCGGG CCGTCTCGG CCGACAGAT AAACGATTTT GCGCCTTACG GGGCGACAA CAGGAACTRA CCTGAATTCG CCTCGAGTT
TBEV signal
-----
deleted C
prM Hypr
-----
· Q K K R G G T D W M S W L L V I G M L G M T I A A T V R K E R D G
201 ACAARAAGAA CCGGGGGGAA CAGACTGGAT GAGCTGGCTG CTGCTARTCG CCAATGCTGGG CATGACAAATC GCAGCTACGG TTCCGAAGGA AAGAGACGGC
TGTTTTCTTT GCCCCCCCTT GTCTGACCTA CTCGACCGAC GAGCATTAGC CGTACGACCC GTACTGTTAG CGTCGATGCC AAGCGTTCCT TTCTCTGCCG
prM Hypr
-----
S T V I R A E G K D A A T Q V R V E N G T C V I L A T D M G S W C D ·
301 AGTACGGTCA TACCGCGGA AGGTAAGGAT GCGCTACCC AAGTGAGAGT GGAAAATGGT ACCTGCGTCA TTCTGCCCAC CGACATGGGC TCTTGGTGTG

```

```

TCATGCCAGT ATGGCGCCT TCCATTCCCTA CGGGGATGGG TTCACTCTCA CCTTTTACCA TGGACGCAGT AAGACCGGTG GCTGTACCCG AGAACCCAC
prM Hypr
401 . D S L S Y E C V T I D Q G E E P V D V D C F C R N V D G V Y L E Y
ATGATAGCCT TTCTTATCAG TGCCTAACCA TAGATCAAGG TGACGAACCT GTTGACGTTG AITGCTTCTG CCGAAACCTG GATGGGGTGT ATCTCGAATA
TACTATCGGA AAGAATACTC ACGCATTGGT ATCTAGTTCC ACTCCTTGGG CAACCTGCAAC TAACGAAGAC GGCTTTGCAC CTACCCCCACA TAGAGCTTAT
prM Hypr
501 . G R C G K Q E G S R T R R S V L I P S H A Q G E L T G R G H K W L
TGGACGGTGT GGTAAACAAG AAGGAAGCAG AACCGAAGCG TCAGTGGCTTA TACCCCTCCA CGCTCRAGGA GAGCTGACCG GACGGGGACA TAAATGGTTG
ACCTGCAGCA CCAATTTGTC TTCCTTCTC TTGGTCTGGC AGTCACGAT ATGGGAGGGT GCGAGTTCTT CTGACTGGC CTGCCCTGT ATTTACCAAC
prM Hypr
601 . E G D S L R T H L T R V E G W V W K N R L L A L A M V T V V W L T L
GAGGGGACT CACTCCGAAC ACATTTGACC CGCTCGAGG GCTGGGCTG SAATAATCGG CTGTTGGCC TCGCTATGGT GACAGTCGT TGGCTCAGC
CTCCCGTGA GTGAGGCTTG TGTAACTGG CCGCAGCTCC CGACCCAGAC CTFTTTAGCC GACAACCGGG AGCGATACCA CTGTCAGCAA ACCGAGTGG
E Hypr
prM Hypr
701 . E S V V T R V A V L V V L L C L A P V Y A S R C T H L E N R D F V
TGGAGTCTGT GGTACTCAG GTGGCAGTGC TGGTGGTCTT CCTCTGTCTT GCCCCTGTCT ACCTGTCAG GTGTACTCAT TTGGAAACA GAGATTTTGT
ACCTCAGCA CCAATGAGCG CACCGTCAGC ACCACCAGA GGAGACAGAA CCGGGACAGA TCGCAGGTC CACATGAGTA AACCTTTGT CTCTAAACA
E Hypr
801 . T G T Q G T T R V T L V L E L G G C V T I T A E G K P S M D V W L
CACCGGCACC CAGGGGACA CTCGGGTAA CCTGGTCTT GACTGGGTG GTTGGTTAC TATTCCCGT GAGGGCAAC CCIATGGA TGTGTGGCTG
GTGGCCGTGG GTCCCTCTT GAGCCATTG GGACCACAA CTGACCCAC CAACGCAATG ATATGGCGA CTCCCGTTT GGAGATACCT ACACCCGAC
E Hypr
901 . D A I Y Q E N P A Q T R E Y C L H A K L S D T K V A A R C P T M G P
GATGCAATCT ATCAGGAGAA TCCCGACAA ACCAGGGAAT ATTECCITCA CCCAAAGCTG TCCGATACAA AGGTCCGCGC TAGGTGCCA ACAATGGAC
CTAGCTTAGA TAGTCTCTT AGGGCTGTG TGGTCCCTTA TAACGGAAT GCCTTTGCAC AGCTATGTT TCCAGCGCCG ATCCACGGGT TGTACCTCTG
E Hypr
1001 . A T L A E E H Q G G T V C K R D Q S D R G W G N H C G L F G K G S
CGGCCACCTT GCGGAGGAA CATCAGGGAG GTACAGTGTG CAAACGGGAC CAGAGTATA GAGGCTGGG TAATCACTGC GGCTGTTCG GCAAGGAAG
GCCGTTGGGA CCGCTCCTT GTAGTCCCTC CATGTACAC GTTTGCCTG GTCTACTAT CTCGACCC ATTAGTGAGC CCGGACAAGC CGTTTCTTC
E Hypr
1101 . I V A C V K A A C E A K K K A T G G E V Y D A N K I V Y T V K V E P
TATTGTCTGT TCGTCAAGG CAGCCTGTGA GGCCAAAAG AAGGCTACTG GGCACGTCTA TGAGCCAAC AAGATCGTTT ATACAGTAA AGTGAACCA
ATACAGCGA ACGCAGTTC GTCCGACCT CCGTTTTTC TTCCGATGAC CCGTCCAGT ACTGGGTTG TTCTAGCAA TATGTCACT TCACCTTGT
E Hypr
1201 . H T G D Y V A A N E T H S G R K T A S F T V S S E R T I L T M G E Y
CACACGGGG ATTACGTGC GCCCAAGAG ACTCATTCG GTCCAAAAC GGCAGCTTC ACCGTCTAT CCGAAAAGAC CATCCTCACT ATGGGGGAGT
GTGTGTCCC TAATGCAGC CCGTTGCTC TGAGTAAGGC CAGCTTTTG CCGTCCAGG TCGCACAGTA GGCTTTTCTG GTAGGAGTA TACCCCTCA
E Hypr
1301 . G D V S L L C R V A S G V D L A Q T V I L E L D K T V E H L P T A
ATGGCGAGT TTCTCTGCT TCCGGGTGG CTAGCGGAGT CGACCTGGCC CAGACAGTCA TCCGGAACG GGATAAACA GTTGAAGATC TGCTACCCG
TACCGCTGA AAGAGACAG ACGGCCACC GATCGCCTCA GCTGACCCG GTCTGTCACT AGGACCTGA CCTATTTGT CACTCGTAG ACGGATGGC
E Hypr

```

```

.....
1401  · W Q V H R D W F N D L A L P W K H E G A R N W N N A E R L V E F G
      TTGGCAGGTG CACAGGGATT GGTTTAACGA CCTTGGCCCTG CCATGGAAAC ATGAAGGAGC GAGAAACTGG AATAATGCAG AGCGACTCGT AGAATTCGGT
      AACCGTCCAC GTGTCCCTAA CCAAAATGCT GGAACGGGAC GGTACCTTTG TACTTCCTCG CTCTTTGACC TTATTACGTC TCGCTGAGCA TCTTAAGCCA
      E Hypr
.....
1501  · A P H A V K M D V Y N L G D Q T G V L L K A L A G V P V A H I E G T
      GCCCTCATG CCGTGAAGAT GGACGTCTAC AATTTGGGTG ATCAGACCGG CTTTCTCCTT AAGGCTCTCG CTGGCGTACC AGTTGCCCCAC ATCGAAGGAA
      CGGGGAGTAC GGCACCTCTA CCTCCAGATG TTAGACCCAC TAGTCTGGCC GCAGAGGGAA TTTCGAGAGC GACCCGATGG TCAACGGGTG TAGCTTCCTT
      E Hypr
.....
1601  · K Y H L K S G H V T C E V G L E K L K M K G L T Y T M C D K T K F
      CGAAGTACCA CCTGAAGTCA GGCCATGTAA CTTCGAGGTG GGGCCTGGAG AAGTTGAAAA TGAAGGTCT TACGTACACA ATGTGTGACA AGACCAAGTT
      GCTTCATGGT GGACTTCAGT CCGGTACAIT GAACGCTCCA CCGGACCTC TCAACTTTT ACTTTCAGAG ATGCATGTGT TACACACTGT TCTGTTCAA
      E Hypr
.....
1701  · T W K R A P T D S G H D T V V M E V T F S G T K P C R I P V R A V
      CACATGGAAG AGGGCCCCCA CAGATAGCGG CCACGATACT GTGGTATGG AGGTGACCTT TTCTGGAACA AAACCTGCA GAATACCCGT GCGGGCTGTA
      GTTACCTTC TCCCGGGGCT GTCTATCGCC GGTGCTATGA CACCCTACCT TCCACTGGAA AAGACCTTGT TTTGGGAGCT CTTATGGGCA CGCCCGACAT
      E Hypr
.....
1801  · A H G S P D V N V A M L I T P N P T I E N N G G G F I E M Q L P P G
      GCTCACGGAT CTCCCGATGT CAATGTTGCT ATGCTGATTA CACCTAACCC TACCATCGAG AATAACCGTG GTGGTTTTAT TGAGATCGAG CTTCCGCCAG
      CGAGTGCCCA GAGGGCTACA GTTACAACA TACGACTAAT GTGGATTGGG ATGGTAGCTC TTATTGCCAC CACCAAAATA ACTCTACGTC GAAGGGGGTC
      E Hypr
.....
1901  · D N I I Y V G E L S Y Q W F Q K G S S I G R V F Q K T K K G I E R
      GCGATAACAT CATCTACGTG GCGGAACCTT CTTACCAGTG GTTTCAGAAA GGGAGTTCAA TTGGGCGGGT CTTCACAAA ACGAAGAAGG GAATCGAAGC
      CGCTATTGTA GTAGATGCAC CCGCTTGAGA GAATGGTCAC CAAAGTCTTT CCTTCAAAT AACCAGCCCA GAAGGTTTTT TGCTTCTTCC CTTAGCTTCC
      E Hypr
.....
2001  · L T V I G E H A W D F G S A G G F L S S I G K A L H T V L G G A F
      ATTGACGGTT ATCGGCGAGC ACGCATGGGA ITTTGGTCC GCAGGGGGAT TCCTGTCTTC TATTGGTAAG GCACCTGCATA CCGTGTGGG GGGCGCATTC
      TAACTGCCAA TAGCCGCTCG TCGTACCCT AAAACCAAGG CGTCCCTTA AGGACAGAAG ATACCATTTC CBTGACGTAT GGCACGACCC CCGCGTAAAG
      E Hypr
.....
2101  · N S I F G G V G F L P K L L L G V A L A W L G L N M R N P T M S H S
      AATTCTATT TTGGGGCGT GGGTTCCTG CCTAAACTCC TGCTGGGAGT AGCCCTGGCC TGCTTGGGAC TGAATATGCG GAATCCGAGC ATGTCATGT
      TTAAGATAAA AGCCCGGCA CCCCAGGAC GGATTTGAGG ACGACCTCA TCGGGACCGG ACCAACCTGT ACTTATAGC CTTAGGCTGC TACAGGTACA
      E Hypr
.....
.....
WNV NS1 protein
.....
2201  · F L L A G V I V L A M T L G V G A D T G C A I D I S R Q
      CATTCCTCTT GGCCGGCGTG CTTGTACTGG CCATGACACT GGGCGTTGGC GCGGACACTG GGTGTGCCAT AGACATCAGC CGGCAA
      GTAAGGAGAA CCGGCCGCAC GAACATGACC GGTACTGTGA CCGCAACCG CCGGTGTGAC CCACACCGTA TCTGTAGTCG GCCGTT

```

PIV-WN/TBEV Hypr with WNV signal (p40)

deleted C
~~~~~

5' UTR

-----

1 AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTTT GTGAGGATTA ACAACAATTA ACACAGTGGC AGCTGTTTCT TAGCACGAAG ATCTCGATGT M S  
TCATCAAGGG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCATAT TGTGTTAAT TGTGTCACGC TCGACAAGA ATCGTGCTTC TAGAGCTACA  
WNV deleted C

-----

101 · K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K R S S K ·  
CTAAGAAACC AGGAGGGGCC GGCAAGAGCC GGGCTGTCTA TTTGCTAATA CGCGGAATGC CCGCGGTGT GTCTTGATT GGACTTAAGC GGAGCTCCAA  
GATTCTTTGG TCCTCCCGGG CCGTTCCTGG CCGACAGAT AAACGATTTT GCGCCTTACG GGGCGCACAA CAGGAACATA CCTCAATTCC CCTCGAGGTT  
WNV signal

-----

WNV deleted C prM Hypr

-----

201 · Q K K R G G K T G I A V M I G M L A C V G A A T V R K E R D G S T ·  
GCAAAAGAAA CGCGGGGAA AGACAGGCAT ACCTGTGATG ATAGGCATGC TGGCTTGTGT CCGAGCAGCT ACCGTGGAA ARGACGGGA CCGAAGCACC  
CGTTTTCTTT GCGCCCCCTT TCTGTCCGTA TCGACACTAC TATCCGTACG ACCGAACACA GCCTCGTCA TGGCAGCCTT TTCTTGGCCT GCCTTCGTGG  
prM Hypr

-----

301 V I R A E G K D A A T Q V R V E N G T C V I L A T D M G S W C D D S ·  
GTGATAAGGG CTGAGGGTAA GGATGCGGCT ACCGAGGTGA GAGTAGAGAA TGGCACTTGC GTAATACTCG CGACTGATAT GGGATCCCTGG TGTGACGATA  
CACTATTCCC GACTCCCAT CCTACGCCGA TGGCTCCACT CTCATCTCTT ACCGTGAACG CATTATGAGC GCTGACTATA CCTAGGACC ACACGTCTAT  
prM Hypr

-----

401 · L S Y E C V T I D Q G E E P V D V D C F C R N V D G V Y L E Y G R ·  
GCCTCAGTIA TGAATGCCGA ACAATAGACC AGGGCCGAAGA ACCTGTGGAC GTTACTGTT TCTGTAGAAA TG'GGATGGC GTTTATCTGG AGTACGGCG  
CGGAGTCAAT ACTTACGCAT TGTATCTCG TCCCGCTTCT TGGACACCTG CAACTGACAA AGACATCTTT ACACCTACCG CAATAGACC TCATGCCGGC  
prM Hypr

-----

501 · C G K Q E G S R T R R S V L I P S H A O G E L T G R G H K W L E G ·  
CTGTGGAAA CAGGAGGCT CAGAACTCG AAGATCTTGT CTGATTCCAA GTCACGGCA ACCAGACTG ACCGGTAGAG GGCACAAGTG GCTTGAAGGG  
GACACCTTTT GTCTCCCGA GTGCTTGAGC TTCTAGACAC GACTAAGTT CAGTGGCGT TCCTTCAAC TGGCCATCT CCGTGTTCAC CGAACTTCCC  
prM Hypr

-----

601 D S L R T K L T R V E G W V W K N R L L A L A M V T V V W L T L E S ·  
GACTCATTGA GGACCCACT GACTAGGGTG GAGGGTTGG TTTGGAAGAA TCGGTTGCTC GCGCTCGTA TGGTACCGT CGTGTGGCTG ACACIGGAGA  
CTGASTAACT CCTGGGTGA CTGATCCAC CTCCCAACC AACCTTCTT AGCCAACGAG CCGGAGCGAT ACCAGTGGCA GCACACCGCA TGTGACCTT  
E Hypr

prM Hypr

-----

701 · V V T R V A V L V V L L C L A P V Y A S R C T H L E N R D F V T G ·  
GTGTCGAC TCGGTTGCT GTGTTGGTT TCCTCTCTG TTTGCCCCA GTTACGGT CCAGGTGAC TCATTTGGA AACAGAGATT TTGTCACCG  
CACACCCTG AGCCCAACA CACAACCAAC AGGAGGAGC AAACCGGGT CACATGGCA GGTCCACATG AGTAAACCTT TTGCTCTAA AACAGTGGC  
E Hypr

-----

· T Q G T T R V T L V L E L G G C V T I T A E G K P S M D V W L D A  
 801 CACCCAGGG AGGACTCGG TAACCCCTGGT GCTTGAACCT GGTGGTTCG TTAATTTAC CGCTGAGGGC AAACCCCTTA TGGATGTTG GCTGGATGCA  
 GTGGGTCCCC TGCTGAGCCC ATTGGGACCA CGAACTTGAC CCACCAACGC AATGATAATG GCGACTCCCG TITGGGAGAT ACCIACACAC CGACCTACGT  
 E Hypr  
 -----  
 I Y Q E N P A Q T R E Y C L H A K L S D T K V A A R C P T M G P A T  
 901 ATCTATCAGG AGAATCCCG ACAAAACCAGG GAATATTGCC TTCACGCCAA GCTGTCCGAT ACAAGGTCG CGCTAGGTG CCCAACAAAG GGACCGGCCA  
 TAGATAGTCC TCTTAGGGG TGTTTGGTCC CTTATAACGG AAGTGGCTT CGACAGGCTA TGTITCCAGC GCGATCCAC GGGTGTGTAO CCTGGCCGGT  
 E Hypr  
 -----  
 · L A E E H Q G G T V C K R D Q S D R G W G N H C G L F G K G S I V  
 1001 CCCTGGCGGA GGAACATCAG GGAGGTACAG TGTGCAACG GGACCAGAGT GATAGAGGCT GGGGTAATCA CTGCGGCCCT TTCGGCAAAG GAATATTGT  
 GGGACCCCT CCTGTAGTC CCTCCATGTC ACACGTTTC CTTGCTCA CTATCTCCGA CCCCATTAGT GACGCGGGAC AAGCCGTTTC CTTATAACA  
 E Hypr  
 -----  
 · A C V K A A C E A K K K A T G H V Y D A N K I V Y T V K V E P H T  
 1101 CGCTTGGCIG AAGGCAGCCT GTGAGGCCAA AAAGAAGGCT ACTGGGCACG TCTATGACGC CAACAAGATC GTTITATACG TGAAAGTGA ACCACACACA  
 GCGAACCGCA TTCCTCGCA CACTCCGGTT TTCTTCCGA TGACCCGTGC AGATACTGCG GTTGTCTAG CAAATATGTC ACTTTCACCT TGGTGTGTGT  
 E Hypr  
 -----  
 G D Y V A A N E T H S G R K T A S F T V S S E K T I L T M G E Y G T D  
 1201 GGGGATPACG TGCGGGCCAA CGGACTCAT TCCGCTCGCA AAACCGCCAG CTTACCGTG TCATCCGAAA AGACATCCT CACTATGGG GAGTATGGC  
 CCCATATGC ACCCCGGT GCTCTGAGTA ACGCCAGCGT TTTCCGGTC GAATGGCAC AGTAGGCTTT TCTGTAGGA GTGATACCC CTTATACCGC  
 E Hypr  
 -----  
 · V S L L C R V A S G V D L A Q T V I L E L D K T V E H L P T A W Q  
 1301 ACGTTCCTCT GCTCTGCCGG GTGGCTAGCG GACTCGACCT GGCCAGACA GTCATCCCTG AACTGGATA AACAGTTGAG CATCTGCCA CCGCTTGGCA  
 TGCAAAAGAGA CGAGACGGCC CACCGATCC CACAGCTGGA CCGGGCTGT CAGTAGGACC TTGACCTATT TTGCAACTC GTAGACGGAT GCGAACCGT  
 E Hypr  
 -----  
 · V H R D W F N D L A L P W K H E G A R N W N N A E R L V E F G A P  
 1401 GGTGCACAGG GATTGGTTA ACGACCTGC CCTGCCATGG AAACATGAAG GAGCGAGAAA CTGGATAAT GCAGAGCGAC TCGTAGAAT CCGTGGCCCT  
 CCAGGTGTC CTAACCAAT TCGTGACG GGACGGTACC TTTTACTTC CTGCTCTTT GACCTTATA CCTCTCGTG AGCATCTTA GCCACGGGA  
 E Hypr  
 -----  
 H A V K M D V Y N L G D Q T G V L L K A L A G V P V A H I E G T K Y  
 1501 CATGCCGTA AGATGGACGT CTACATCTG GGTGATCAGA CCGCGCTCT CTTAAAGCT CTGCTGGCG TACCAGTGC CCACATCGAA GGAACGAGT  
 GTACGGCACT TCTACCTGA GATGTTAGAC CCACTAGTCT GGCCCAAGA GGAATTTGA GAGCGACCG ATGTTCAAC GGTGTAGCT CTTGCTTCA  
 E Hypr  
 -----  
 · H L K S G H V T C E V G L E K L K M K G L T Y T M C D K T K F T W  
 1601 ACCACCTGAA GTCAGGCCAT GTAACCTCG AGTGGGCCCT GGAGAAGTTG AAAATGAAAG GTCTIACGTA CACRATGTGT CACAAGACCA AGTTCACATG  
 TGGTGGACT CAGTCCGGA CATTGAACGC TCCACCCGGA CCTCTTCACT TTTTACTTTC CAGAATGCA GTGTIACACA CTGTCTGGT TCAAGTGTAC  
 E Hypr  
 -----  
 · K R A P T D S G H D T V V M E V T F S G T K P C R I P V R A V A H  
 1701 GAAGAGGCC CCCACAGTA CCGGCCACGA TACTGTGCT ATGGAGTGA CCTTTCTG AACAACCC TCGAATAC CCGTCCGGC TGTAGTCCAC  
 CTCTCCCGG GGGTCTAT CGCCGTGCT ATGACACCAC TACCTCCACT GGAARAGACC TTGTTTGGG ACGTCTATG GGCACCGCC ACATCGAGTG  
 E Hypr  
 -----  
 G S P D V N V A M L I T E N P T I E N N G C G F I E M Q L P P G D N  
 1801 GGATCTCCCG ATGTCATGT TGCTATGCTG ATTACACCTA ACCCTACCAT CGAGAATAAC GGTGGTGGT TTATTGAGAT CGAGCTTCCG CCAGGGCGATA  
 CCTAGAGGCC TACAGTTACA ACGATACGAC TAATGTGAT TGGGATGGA GCTCTTATG CCACCACCA AATAACTCTA CGTCCAGGC GGTCCGCTAT  
 E Hypr

```

.....
1901  · I I Y V G E L S Y Q W F Q K G S S I G R V F Q K T K K G I E R L T ·
ACATCATCTA CGTGGCGAA CTCTCTTACC AGTGGTTCA GAAAGGAGT TCAATGGGC GGTCTTCCA AAAAACGAAG AAGGGAATCG AACGATTGAC
TGTAGTAGAT GCACCCGCTT CAGAGAATGG TCACCAAGT CTTCCCTCA AGTTAACCG CCCAGAAGT TTTTGCCTC TTCCTTAGC TTGCTAATG
E Hypr
.....
2001  · V I G E H A W D F G S A G G F L S S I G K A L H T V L G S A F N S ·
GGTATCGGC GAGCAGCAT GGGATTTGG TTCGCGGGG GATTCCTGT CTCTATTGG TAAGGCATG CATACCGTGC TGGGGGGCC ATTCAATTCT
CCATAGCCG CTCGTGCGTA CCCPAAACC AAGGCGTCCC CCAAGGACA GAAGATAACC ATTCCGTGAC GTATGGCAG ACCCCCCGG TAAGTTAAGA
E Hypr
.....
2101  · I F G G V G F L F K L L L G V A L A W L G L N M R N P T M S M S F L ·
ATTTTCSSGG GCCTGGGTT CCGCCTAAA CTCCTGCTGG GAGTAGCCT GGCCTGGTGG GACTGAATA TCGGGAATCC GACGATGCC ATGTCATTCC
TAAAGCCCC CGCACCCAA GGACGGATTT GAGGACGACC CTCATCGGA CCGGACCAAC CCTGACTTAT ACGCCTTAGC CTGCTACAGG TACAGTAAG
E Hypr
.....
WNV NS1 protein
.....
2201  · L A G V L V L A M T L G V G A D T G C A I D I S R Q ·
TCTTGGCCGG CGTGTGTA CTGGCATGA CACTGGCGGT TGGCGCGAC ACTGGGTGTG CCATAGACAT CAGCCGGCAA
AGAACCGGCC GCACGAACAT GACCGTACT GTGACCCCA ACCGCGGCTG TGACCCACAC GGTATCTGTA GTCGCCGTT

```

# Sequence Appendix 4. WN PIV constructs expressing rabies virus G protein.

## WN ( $\Delta$ CprME)-Rabies PIV sequence (partial)

N-terminus of C

5' UTR

M S

1 AGTAGTGGC CTETGTGAGC TGACAAACTT AGTAGTGTCT GTGAGGATTA ACAACAATTA ACACAGTGGC AGCTGTTTCT TAGCACGAAG ATCTCGATGT  
TCATCAAGGG GACACACTCG AC<sup>1</sup>GT<sup>1</sup>TGAA TCATCACAAA CACTCCTAAT TGGTGTAAAT TGTCACCCG TCACACAAAAG ATCTGTCTTC TAGAGCTACA  
N-terminus of C

· K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R ·

101 CTAGAAACC AGGAGGGCCC GGCAAGAGCC GGGCTGTCTA TTTGCTAATA CGCGGAATGC CCCGGTGT<sup>1</sup> GTCTTTGATT GGACTTAAGC AANAGAAGCG  
GATTC<sup>1</sup>TTGG TCCTCCCGCG CCGTTC<sup>1</sup>CGG CCGACAGAT AAACGATTTT GCGCCTTAGG GGGCGCACAA CAGGA<sup>1</sup>ACTAA CTTGAATT<sup>1</sup>CG TTTTCTTC<sup>1</sup>CG  
N-terminus of C Rabies-G signal

partial C signal Rabies-G protein

· G G K T G I A V I V P Q A L L F V P L L V F P L C F G K F P I Y T ·

201 AGGGGGCAAG ACTGGPATACT CTCTGATCCT TCCTCAGGCT CTTTGTGTTG TACCCCTT<sup>1</sup>CT GGTATTTC<sup>1</sup>CC CTTTGTGTTG GTA<sup>1</sup>AATTC<sup>1</sup>CC TATCTATAC<sup>1</sup>C  
TCCCCCGTTC TGACCATATC GACACTAGCA AGGAGTCCGA GAAAACAAC ATGGGAACGA CCATAAACCG CAACCA<sup>1</sup>AAAC CATTAAAGG ATAGATATGG  
Rabies-G protein

· I P D K L G P W S P I D I H H L S C P N N L V V E D E G C T N L S G ·

301 ATCCCTGATA AGCTCGGGCC TTGGAGTCCC ATTGATAATC ACCATT<sup>1</sup>AGG CTGCCCA<sup>1</sup>AAAC AACCTCGT<sup>1</sup>CG TTGAGGATGA AGGGTGC<sup>1</sup>ACT AATCTTTC<sup>1</sup>CG  
TAGGGACTAT<sup>1</sup> TCGAGCCCGG AACCTCAG<sup>1</sup>GS TAACATAA<sup>1</sup>AG TGGTAACTC GACGGGT<sup>1</sup>TTG TTGGAGCAGC AACTCCTACT TCCCAC<sup>1</sup>TGA ATAGA<sup>1</sup>AAGAC  
Rabies-G protein

· F S Y M E L K V G Y I S A I K M N G F T C T G V V T E A E T Y T N ·

401 GATTTCTGTA CATGGAGITG AAGTGGGCT ATATTTCAGC CATTAAGATG AACGGCTT<sup>1</sup>TA CTGTACAGG AGTCGTG<sup>1</sup>ACC GAAGCCGAGA CATATACAA  
CTAAAGATG<sup>1</sup> TACCTCAAC TTTCA<sup>1</sup>CCGA TATAAAGT<sup>1</sup>CG GTAATCT<sup>1</sup>AC TTGCCGAAT GAACATG<sup>1</sup>TCC TCAGCACTGG CTTCCGCT<sup>1</sup>CT GTATATGTT<sup>1</sup>  
Rabies-G protein

· F V G Y V T T T F K R K H F R P T P D A C R A A Y N W K M A G D P ·

501 TTTCTGGGA TACGTACCA CCACCTCAA GAGAAACAC TTCGGCCAA CGCTGAGCC TTGTCGGCC GCTTACA<sup>1</sup>ACT GGAAGTGGC AGGAGATCT  
AAGCACT<sup>1</sup>AT TCGACTG<sup>1</sup>GT GGTGAAGT<sup>1</sup>CTCTTT<sup>1</sup>TGTG AAGCCGGGT<sup>1</sup>CG GCGGACTG<sup>1</sup>CG AACAGCCCGG CGAATG<sup>1</sup>TGA CTTCTAC<sup>1</sup>CG CCTCTAGGA  
Rabies-G protein

· R Y E E S L H N P Y P D Y H W L R T V K T T K E S L V I T S P S V A ·

601 CGATATGAG AATCTGCA CAACCCGTAT CTTGATTACC ATTGGCTCCG QACACTCAAG ACTACCAAGG AGACTGTGGT CATTATATCA CCAAGCGTGG  
GCTACTACTT<sup>1</sup> TTAGAGACGT GTTGGGATA GACTAATGG TAAGCGAGCC CTCTCAGT<sup>1</sup>TC TGATGGT<sup>1</sup>TCC TCTCAGACA GTAATATAGT GTTCCG<sup>1</sup>CACC  
Rabies-G protein

· D L D P Y D R S L H S R V F P G G N C S G V A V S S T Y C S T N H ·

701 CCGATCTCA TCCTTATGAT AGATCCCTGC ACAGTAGGGT TTTTCTGGC GGGAAITGTA GCGGTGTGG AGTATCAAGT ACCTACTGCT CCACTAACCA  
GCTAGAACT AGGAATAC<sup>1</sup>TA TCTAGGGAGC TGTCATCCCA AANAGGACCG CCGTTA<sup>1</sup>ACAT CGCCACAGC TCATAGTCA TGGATGACA GGTGATG<sup>1</sup>GT  
Rabies-G protein

· D Y T I W M P E N P R L G M S C D I E T N S R G K R A S K G S E T ·

801 CGACTACTAT ATATGGATGC CTGGAACCC TCGACTCGGT ATGAGTGGG ACATTTT<sup>1</sup>TAC GAAC<sup>1</sup>CAAGG GGAAGCGGG CAICTAAGGG GTCTGAACA  
GCTGATGTA TATACCTAGC GACTCTGGG AGCTGAGCCA TACTCAAGC TGTA<sup>1</sup>AAAATG CTTGAGTGGC CCGTTCGCCC GTAGATTC<sup>1</sup>CC CAGACTT<sup>1</sup>GT  
Rabies-G protein

· C G F V D E R G L Y K S L K G A C K L K L C G V L G L R I M D G T W ·

901 TCGGGTITG TTGATGAGCG GGGTITGTAT AAATCTCTTA AAGCCCTC<sup>1</sup>TA AACCTGAAA CTCTGTGGC TACTGGGGCT GCGCCTGATG GACGGACAT  
ACGCCAAAC AACTACTG<sup>1</sup>CC CCCCACATA TTTAGAGAA<sup>1</sup>T TCCCGCGGAC AITCGACT<sup>1</sup>TT GAGACACCG ATGACCCCGA CCGGACTAC CTGCCCTGTA  
Rabies-G protein

· V A M Q T S N F T K W C P P G Q L V N L H D F R S D E I E H L V V ·

1001 GGTGGCTAT GCAGACAGC ATGAAACAA AGTGGTCTCC CCTCTGTCAG CTGTTAA<sup>1</sup>TG TCACAGACT TAGGTCTGAC GAAATCGAG ACCTGTGCT  
CCACCGATA CGTCTGT<sup>1</sup>CC TTACTTT<sup>1</sup>CT TCACCACAG GGGACAGT<sup>1</sup>AG ACCCAATTAG AGTGTCTGA ATCCAGACT<sup>1</sup>CT CTTTAGCT<sup>1</sup>CG TGAACACCA  
Rabies-G protein

```

.....
. E E L V K K R E E C L D A L E S I M I T K S V S F R R L S H L R K
1101 GGAGAACG GTGAGAAC GCGAGAGTG CCTGGAGCA CTTCAGATA TTATGACCAC CAAATCCGTT TCCTTCAGAA GACTGAGCCA CTGCGAAG
CCTCCTGAC CACTTCTTG CGCTTCTCAC GCACCTGGCT GAATCTCTAT AATACTGGTG GTTTAGCCAA AGGAGTCTT CTGACTCGT GGACGTTT
Rabies-G protein
.....
. L V P G F G K A Y T I F N K T L M E A D A H Y K S V R T W N E I I P
1201 CTGGTCCAG GGTTCGGAA GGTATTACT ATTTTCAACA AGACTCTTAT GGAGCGGAT GCCCTTATA AGTCAGTTAG GACTTGGAT GAGATAATC
GACCACGTC CCAAGCCCTT CCGAATAWGA TAAAAGTTGT TCTGAGAATA CCTCCGCTA CGGGTAATAT TCAGTCAATC CTGAGCCTTA CTCTATTAG
Rabies-G protein
.....
. S K G C L R V G R C H P H V N G V F P N G I I L S P D G N V L I
1301 CCTCAAAGG ATGCTGAGA GTCGTTGGA GATGCCCC CCATGTCAT GGGTGTCT TTAACGAT CTTCTGGA CCTGACGGA ACGTGTGAT
GGAGTTCCT TACAGACTCT CAGCCACCT CTACGGTGG GGTACATTA CCCCACAAGA AATTGCCTTA GTAGGACCTT GACTGGCTT CCCCAGTA
Rabies-G protein
.....
. P E M Q S S L L Q Q H M E L L V S S V I P L M H P L A D P S T V F
1401 TCCGAGATG CAATCTCC TCTGACGA ACACATGGA CTCCGTGTG CTTCAGTAT ACCCTGATG CACCCACTG CCGACCCAG CACTGTGTC
AGGCTCTAC GTTAGAGGG AAGACCTCT TGTGTACCTT GAGGACACA GAATCACTA TGGGACTAC GTGGTGGAC GCGTGGGCTT GTGACACA
Rabies-G protein
.....
. K N G D E A E D F V E V H L P D V H E R I S G V D L G I P N W G K Y
1501 AATATCGG ATGAGCGGA AGACTTTGG GAATTCACC TCCCGATGT ACACGAAAG ATATCTGAG TAGACCTGG CCTTCTAAT TGGGTAAAT
TTTTACCG TACTCCCTCT TCTCAACAC CTTCAAGTG ACGGGCTCA TGTGCTTCC TATAGACTC ATCTGGACC GGAAGGATTA ACCCAACTA
Rabies-G protein
.....
. V L L S A G A L T A L M L I I F L M T C W R R V N R S E P T Q H N
1601 ACGTCTCT GAGTCCGGT GCCTGACCG CTTGATGCT GATCATTTT CTGATGACT GCTGGCCGAG GGTGAATCG TCCGAGCGA CACAGACAA
TGCAGGAGA CTCAGCCCA CGGAAGTGC GAAACTACA CTACTAARA GACTACTGA CGACCGCTC CCACTTAGG AGGCTCGCTT GTGCTGTGT
Rabies-G protein
.....
FMDV 2A
.....
. L R G T G R E V S V T P Q S G K I I S S W E S Y K S G E E T G L N
1701 TCTCAGAGG ACAGCCGGG AGTARSPT GACTCCGAA TCTGGCAGA TTATTACTAG TTGGAGAGT TACAGTCTG GAGGAGAGC TGCTTCAAT
AGACTCTCC TGCCTGGCC TCTATTACA CTGAGGCTT AGACCGTCT AATAATCAT AACCCCTCA APTGTCAGC CTCTCTCTG ACCCAACTA
preNS1 signal
.....
FMDV 2A NS1 signal
.....
. F D L L K L A G D V E S N P G P A R D R S I A L T F L A V G G V L L
1801 TTGACTCTG TCAACTTGC AGCGATCTA GAATCAAATC CTGACCCCG CCGGACAGG TCATAGCTC TCAGCTTCT CCGACTTGA GAGTCTG
AAACTAGAG AGTTTGAAG TCCCTACAT CTAGTTTAG GACTGEGCG GCGCTGTCC AGGTATGAG AGTGCAAGA GCGTCAACT COTCAAGAG
NS1 signal
.....
NS1
.....
. F L S V N V H A D T G C A I D I S R Q E I R C G S G V F I H N D V
1901 TCTCCTCTC CTGACGAGG CAGCTGACA CTGGCTGTC CATAGACAIC ACCCGCAAG ASCTGAGATG TGAAGTGA CTGTTATAC ACAATGATG
AGAGGAGAG GCACCTGCAC GTGGACTGT SACCACACG GTATCTGTAG TCGCCCTTC TCGACTTAC ACCTCACTT CACAGTATG TCTTACTACA

```

WN ( $\Delta$ C)-Rabies G PIV sequence (partial).

5'UTR

-----

of C N-terminus

-----

M S

1 AGTAGTCCG CTGTGTGAG TCACAAACTT AGTAGTGTT GTGAGGATTA ACAACAATTA ACACAGTGGC AGCTGTTCT TAGCACGAG ATCTCGATGT  
 TCATCAAGCG GACACACTCG ACTGTTTGA TCAATCACAA CACTCCATAT TGTGTGTAAT TGTSTCACGC TCGACAAAGA ATCGTGCTTC TAGAGCTACA  
N-terminus of C

-----

· K K P G G P G K S R A V N M L K R G M P R V L S L I G L K Q K K R ·

101 CTAGAAGCC AGGAGGCC CCACAGGCC CGCTGTCAA TPTGCTPAAA CGCGGATGC CCGCGTGTG TCTCTGATP GACTTAAAC AAAAGAGCG  
 GATTCCTTGG TCTTCCCGG CGTTCCTCG CCGACAGTT ATACGATTT GCGCCTTAC GGGCGACAA CAGGAACTA CCTGAATCG TTTTCTTGC  
N-terminus of C Rabies-G protein

-----

partial C signal Rabies-G signal

-----

· G G K T G I A V I V P Q A L L F V P L L V F P L C F S K F P I Y T ·

201 AGGGGGCAAG ACITGGPATG CTGTGATCGT TCCTCAGGCT CTTTCTTTC TACCTTTCCT GGTATTTCCT CTTGCTTTC GTAATTTCG TATCTATACC  
 TCCCCGTTT TGACCATATC GACACTAGCA AGGAGTCCGA GAAACAACAC ATGGAACGA CCTAAAGGG GAACCAAC CATTAAAGG ATAGATATGG  
Rabies-G protein

-----

· I P D K L G P W S P I D I H H L S C P N N L V V R D F G C T N L S C ·

301 ATCCCGATA AGCICGGGCC TTGGATGCC ATTGATATC ACCATTGAG CTGCCAAC ACCCTGCTG TTGAGATGA AGGGTCACT AATCTTCTG  
 TAGGGACTAT TCGAGCCCG ACCCTCAGG TAACATATAG TGGHAACTC CACCGGTTTG TTGGAGCAG AACTCTACT TGGACGTGA TAGAAAAGAC  
Rabies-G protein

-----

· F S Y M E L K V G Y I S A I K M N G F T C T G V V T R A E T Y T N ·

401 GATTTCCATA CATGGAGTTC AAATGCGCT ATATTTCAG CATTAGATG AAGCGTTTA CTGTACAGG AGTCTGACC GAAGCCGAGA CATATACAA  
 CTAAGAGATG GTACCTCAAC TTTCACCGA TATAAAGTC GAAATTCAC TTGCGAAT CAACATGTC TCAGCACTGG CTTGCGTCT GTATATGTT  
Rabies-G protein

-----

· F V G Y V T T T F K R K H F R P T P D A C R A A Y N W K M A G D F ·

501 TTTCGTGGA TACTGACCA CACCTTCAA GACAAAAC TTCCGCCAA CGCTGACG CTTCGSGCC GCTTACACT GGAAGATGC AGGAGTCTT  
 AAGCACCTT ATGCACTGT GGTGAAGT CTCTTTTGT AAGCGGCTT CCGGACTGC AACAGCCCG CGAATGTGA CTTGATATC TCTCTAGGA  
Rabies-G protein

-----

· R Y F E S L H N P Y P D Y H W L R T V K T T K E S L V I I S P S V A ·

601 CGRTGTAAG AATCTTCCA CAACCCATC CTTATTACC ATTCCTCCG GACATCAAG ACTACCAAG AGACTCTGT CATTATATCA CAAGCGTGG  
 GCIATACTTC TTAGAGAGT GTTGGGATA GGACIAATG TAACCGAGC CTGTGATTC TGATGTTCC TCTCAGACA CTAATATAG GTTCTCGACC  
Rabies-G protein

-----

· D L D P Y D R S L H S R V F P G G N C S G V A V S S T Y C S T N H ·

701 CCGATCTGA TCTTATGAT AGATCCCTG ACACTAGGCT TTTTCTGSC GSAATTGA CCGTGTTCG AGTATCAAGT ACCTACTGCT CCATCAACA  
 GGCTAGAACT AGGAATACTA TCTAGGAGC TCTCATCCA AAAAGGAGC CCTTAAAT CCGCACACG TCATAGTTCA TGAATGACA GTGATTTGT  
Rabies-G protein

-----

· D Y T I W M P E N P R L G M S C D I F T N S R G K R A S K G S E T ·

801 CGACTACACT AATGATGTC CTGAGACCC TCGACTCGT ATCAGTTCC ACATTTTAC GAATCAAGG GGCAGCGGG CATCTAAGG GTCTGAACA  
 GCTGATGTA TATACCTAG GACTCTTGG AGCTGAGCA TACTCAAGC TGTAAATG GAGACAGC ATGACCCCG CTTGATGCT CCGTTCGCC CAGACTTTG  
Rabies-G protein

-----

· C G F V D E R G L Y K S L K G A C K L K L C G V L G L R L M D C T W ·

901 TCGGCTTTC TTGATGAGC GGGTGTGAT AAATCTCTA AAGGGCTG TAAGCTGAAA CTCTGTGCG TACTGGGCT GCCTGTGAT GACGGACAT  
 ACGCCAAAC AACTACTCG CCCCACATA TTACAGAAAT TTCCCGGAC ATTGCACTT GAGACAGC ATGACCCCG GCGGACTAC CTGCGGTGA  
Rabies-G protein

-----

· V A M Q T S N E T K W C P P G Q L V N L H D F R S D E I E H L V V ·

1001 GGTGGCTAT GAGACAAGC AATGAACAA AGTGGTCTC CCTGGTCAG CTGGTAAAT TGACAGACT TAGGTCTGAC GAAATCGAGC ACCTTGTGT  
 CCGACGATA GTCCTTTCG TACTTTCTT TCACCACAG GAGCCAGTC GACCAATTAG ACGTGTGAA ATCCAGACT CTTTACTTCG TGGAAACAA  
Rabies-G protein

-----

· E E L V K K R E E C L D A L E S I M T T K S V S F R R L S H L R K ·

1101 GGAGGAATG GTGAGAAC GCGAAGATG CTTGAGGTA TTATGACCAC CAATCGTT TCTTCAGTA GACTGAGCA CCTCCGAAAG  
 CTTCTTAC CACTTCTTG CGTCTCAC GGACTCGT GAATCTCAT AACTACTGT GTTTAGCAA AGGAAGICT CTGACTGCT GGACCTTTC  
Rabies-G protein

-----

L V E G F G K A Y T I F N K T L M E A D A H Y K S V R T W N E I I P .  
1201 CTGTTCCAG GGTTCGGGAA GGCTTATACT ATTTTCAACA AGACTCTTAT GGAGGCGGAT GCCCATATATA AGTCAGTTAG GACTTGGGAT GAGATAATTC  
GACCACGGTC CCAAGCCCTT CCGAATATGA TAAAGTTGT TCTGAGAATA CCTCCGCTA CGGGTAATAT TCAGTCAATC CPGAACCTTA CTCTATTAG  
Rabies-G protein

· S K G C L R V G G R C H P H V N G V F F N G I I I G F D G N V L I ·  
1301 CCTCCAAGG ATGTCGAGA GTGGTGGGA GATGCCACC CCATGTCAT GGGGTGTCT TTAACGGAAT CATCCTGGGA CCTGACGGGA ACCTGCTGAT  
GGAGGTTTC TAGACTCT CAGCCACCCT CTACGGTGG GGTACAGTA CCCCACAAGA AATTCCTTA GTAGGACCTT GGACTGCTCT TCACAGCTA  
Rabies-G protein

· P E M Q S S L L Q Q H M E L L V S S V I P L M H P L A D P S T V F ·  
1401 TCCCGAGATG CAATCTTCCC TTCTGCAGCA ACACATGGAA CTCCTGGTGT CTTCACTGAT ACCCCTGATG CACCCACTGG CCGACCCAG CACTGTGTT  
AGGGCTCTAC GTTAGAAGG AAGACGTCT TGTGTACCTT GAGGACCACA GAAGTCACTA TGGGGACTAC GTGGGTGACC GGCTGGGGTC GTGACACAAG  
Rabies-G protein

K N G D E A E D F V E V H L P D V H E R I S G V D L G L P N W G K Y ·  
1501 AAAATGGCG ATGAGGCGGA AGACTTTGTG GAAGTTCACC TGCCCGATGT ACACGAAAGG ATACTGGAG TAGACTGGG CCTTCTAAT TGGGTAAAT  
TTTTACCGT TACTCCGGCT TCTGAACAC CTTCAGTGG ACGGGCTACA TGTCCTTTC TATAGACCTC ATCTGGACCC GGAAGGATTA ACCCATTA  
Rabies-G protein

· V L L S A G A L E A L M L I I F L M T C W R R V N R S E P T Q H N ·  
1601 ACGTCTCTT GAGTCCGGT CCCTGACCG CTTGATGCT GATCATTTT CTGATGACCT GGTGGCGGAG GGTGAATCG TCCGACCGGA CAGACAGAA  
TGCACGAGA CTCACGCCA CGGAACCTGC GAAACTACGA CTAGTAAAA GACTACTGA CGACCCTTC CCATTAAGG AGGCTGGCT GTGCTGTGT  
Rabies-G protein  
FMDV 2A

· L R G T G R E V S V T P Q S G K I I S S W E S Y K S G G E T G L N ·  
1701 TCTCAGGGG ACAGGCGGG AAGTAAGTGT GACTCCGCA TCTGCCAAGA TTATTACTAG TTCGGAGACT TACAGTCTG GAGGAGAGC TGGSTGAAT  
AGACTCTCC TGTCCGGCC TTCACTACA CTGAGGGCTT AGACCTTCT AATAATCATC AACCTCTCA ATGTTACAG CTCCTCTCTG ACCCAACTTA  
C/prM signal

FMDV 2A  
F D L L K L A G D V E S N P G P G G K T G I A V M T G L I A C V G A ·  
1801 TTTGATCTC TCAAACITGC AGGCGATGA GAATCAAATC CTGGACCCGG AGGAAAGACC GGIATTCAG TCATGATTG CCTGATCCG TCGTAGGAG  
AAACTAGACG AGTTGAACG TCCGCTACAT CTAGTTTAG GACTCTGGCC TCCTTTCTG CCATAACGTC AGTACTAACC GGACTAGCGG ACGCATCTC  
C/prM signal

prM  
· V T L S N P Q G K V M M T V N A T D V T D V I T I P T A A G K N L ·  
1901 CAGTACCCT CTCTAATTC CAAGGGAAGG TGATGATGAC GGTAAATGCT ACTGACGTA CAGATGTCAT CAGCATCCA ACAGCTGGT GAAAGAACCT  
GTCATGGGA GAGATTGAG GTTCCCTTCC ACTACTACTG CCATTAAGTA TACTGCACT GCTACAGTA GTGCTAAGT TGTCCGACG CTTTCTTGA  
prM

· C I V R A M D V G Y M C D D T I T Y E C P V L S A G N D P E D I D ·  
2001 ATGATTTTC AGAGCAATGG ATGTGGATA CATGTGGAT GATACTATCA CTTATGAATG CCCAGTGGT TCGGCTGTA ATGACCGA AGACATGAC  
TAGCTAACG TCTCTTACC TACACCCAT GTACACGCTA CTATGATAGT GAATACTTAC GGGTACGAC AGCCGACCAT TACTAGTCT TCTGTAGCT  
prM

C W C T K S A V Y V R Y G R C T K T R H S R R S R R S L T V Q T H G ·  
2101 TGTGTTGCA CAAGTCAGC AGTCTACGTC AGGTATGAA GATGACCAA GACAGCCAC TCAAGACCA GTCCGAGGTC ACTGACACTG CAGACACAG  
ACAACCACGT GTTTCAGTGG TCAGATGACG TCCATACCTT CTACTGTGTT CTGTCCGGT AGTTCCTGCT CAGCTCCAG TGACTGTAC CTTGTGTGC  
prM

· E S T L A N K K G A W M D S T K A T R Y L V K T E S W I L R N P G ·  
2201 GAGAAAGC TCTAGCGAAC AAGAAGGGGG CTTGGATGGA GAGCACCAAG GCCACAAGT ATTTGGTAAA AACAGATCA TCGATCTGA GGAACCTGG  
CTCTTCTG AGATCGCTG TTCTTCCCC GAACCTACCT GTCTGTGTT CCGGTTTCCA TAAACCATTT TTGTCTTAGT ACCTAGAAT CCTTGGGAC  
prM

· Y A L V A A V I G W M L G S N T M Q R V V F V V L L L L V A P A Y ·  
2301 ATATGCCCTG GTGGCAGCG TCATTGTTG GATGCTTGG AGCAACCCA TCRAGAGAT GTGTTTCTC GTGCTATTGC TTTTGGTGGC CCCAGCTAC  
TATACGGAC CACCSTGGC AGTAACCAAC CTACGAACCC TCGTTGGTGT AGTCTCTCA ACACAACAG CACGATACG AAAACCAAGG GGTCTGAATG  
E

prM  
S F N C L G M S N R D F L E G V S G A T W V D L V I E C D S C V T ·  
2401 ACCTTAATC CCTTGGAA GAGCAACAGA GACTTCTGG AAGGAGTGC TGGAGCAACA TGGGTGGATT TGGTCTCGA AGGCGACAG TCGTGTACTA  
TCGAATTA CGAACCTTA CTCTTGTCT CTGAAGAACC TTCTCACAG ACCTCGTGT ACCCACCTAA ACCAAGAGCT TCCGCTGTG ACGCACTGAT  
E

· M S K D K P T I D V K M M N M E A A N L A E V R S Y C Y L A T V S ·  
 2501 TCATGTCGTA GGCAAGGCT ACCATCGATG TGAAGATGAT GAATATGGAG GCGGCCAACC TGCCAGAGGT CCGCAGTTAT TGCTATTGGG CTACCGTCAG  
 AGTACAGATT CCGTTCGGG TGGTAGCTAC ACTTCTACTA CTTATACCTC CGCGGTTGG ACCGTCCTCA GCGTCAATA ACGATAAACC GATGGCAGTC  
 E

· D L S T K A A C P A M G E A I I N D K R A D P A F V C R Q G V V D R ·  
 2601 CGATCTCTCC ACCAAAGCTG CGTGCCCGCG CATGGGAGAA GCTCACAAATG ACAACCGTGC TGACCCAGCT TTTGTGTGCA GACAGGAGT GGTGGACAGG  
 GCTAGAGAGG TGGTTTCGAC GCACGGGSCG GTACCCCTCTT CGATGCTTAC TGTTTGCACG ACTGGGTGCA AAACACAGCT CTGTCTCTCA CCACCTGTCC  
 E

· G W G N G C G L F G K G S I D T C A K F A C S T K A I G R T I L K E ·  
 2701 GCCTGGGSCA ACGGCTCGGG ACTATTTGGC AAAGGAAGCA TTGACACATG CGCCAAATTT GCCTGTCTTA CCAAGGCAT AGGAAGAACCC ATTTTGAAG  
 CGACCCCGT TGCCGACGCC TGATAAACCG TTTCTTCTG AACTGTGTAC GCGGTTTAAA CGGACGAGAT GGTTCGTTA TCCTTCTTGG TAAACTTTC  
 E

· N I K Y E V A I F V H G P T T V E S H G N Y S T Q V G A T Q A G R ·  
 2801 AGAATATCAA GTACAGAGTG GCCATTTTGG TCCATGGACC AACTACTGTG GAGTCCGACG GAACTACTC CACACAGGTT GGAGCCACTC AGCAGGGGAG  
 TCTTAPAGTT CATGCTTCAC CGGTAAAAAC AGGTACCTGG TTGATGACAC CTCAGCGTGC CTTTGATGAG GTGTGCCAA CCTCGGTGAG TCCGTCCTC  
 E

· F S I T P A A P S Y T L K L G E Y G E V T V D C E P R S G I D T N ·  
 2901 ATTGAGCACT ACTCTGCGG CGCCTTCATA CACACTAAG CTTGGAGATG ATGGAGAGGT GACAGTGGAC TGTGAACACC GGTGAGGATG TGACACCAAT  
 TAAGCTGATG TGAGGACGCC GCGGAAGTAT GTGTGATTC GAACTCTTCA TACCTCTCCA CTGTACCCG ACACCTGGTG CCAGTCCCTA ACTGTGGTTA  
 E

· A Y Y V M T V G T K T F L V H R E W F M D L N L P W S S A G S T V W ·  
 3001 GCATCTACG TGATGACTGT TGGAACAAAG ACCTTCTTGG TCCATCGTGA STGSTTCATG GACCTCAACC TCCCTTGGAG CAGTGTCTGA ACTACTGTGT  
 CGTATGATG ACTACTGACA ACCTTCTTTC TGCAAGAAC AGGTAGCATC CACCAAGTAC CTGGAGTTGG AGGGAACCTC GTACAGACTC TCATGACACA  
 E

· R N R E T L M E F E E P H A T K Q S V I A L G S Q E G A L H Q A L ·  
 3101 GGAGGAACAG AGACAGGTTA ATGGAGTTTG AGGAACCAACA CGCCACCAAG CAGTCTGTGA TAGCAITGGG CTCACAGAGG GGAGCTCTCG ATCAAGCTTT  
 CCTCTTGTG TCTTGGCAAT TACCTCAAAC TCCTTGTGTG GCGGTCTTTC GTGACACACT ATCGTAACCC GAGTGTCTC CCTCGAGAGC TAGTTCGAAA  
 E

· A G A I P V E F S S N T V K L T S G H L K C R V K M E K L O L K G ·  
 3201 GCCTGGAGCC ATTCCTGTGG AATTTTCAAG CAACACTGTC AAGTTGAGCT CCGGTCAATT GAAGTGTAGA GTGAAGATGG AAAAATTGCA GTTGAAGGGA  
 CCGACCTCGG TAAGGACACC TTAAGAGTTC GTTGTGACAG TTCAACTGTA TGTATCACCA GCGCAGIAAA CTCACATCT CACTTCTACC TTTTAACTC CAACTTCCT  
 E

· T T Y G V C S K A F K F L G T P A D E G H G T V V L E L Q Y T G T D ·  
 3301 ACAACCTATG GCCTCTGTTC AAACGCTTTC AAGTTTCTTG GGACTCCCGC AGACACAGGT CACGGCACTG TGTGTTTGA ATTCAGTAC ACTGGACGGG  
 TGTGGATAC CGCAGACAAG TTTCCGAAAG TCAAGAAGAC CCTGAGGGCG TCTCTCTCCA GTCCGCTGAC ACCACAACCT TAACGTCATG TGACCGTGGC  
 E

· G P C K V P I S S V A S L N D L T P V G H L V T V N P F V S V A T ·  
 3401 ATGACCTTGG CAAGTTCCTT ATCTCGTCAG TGGCTTCATT GAACGACCTA ACGCCAGTGG GCAGATTGGT CACTGTCAAC CCTTTTGTG CAGTGGCCAG  
 TACCTGGACG GTTCAAGGA TAGAGCAGTC ACCGAAGTAA CTGTCTGGAT TGCGGTCAAC CGTCAACCA GTGACAGTGG GGAAAACAAA GTACCGGTG  
 E

· A N A K V L I E L E P P F G D S Y T V V G R G E Q Q I N H H W H K ·  
 3501 GGCACACCT AAGTTCCTGA TTGAATTGGA ACCACCTTTT GGAGACTCAT ACATAGTGGT GGGCAGAGGA GAACACAGA TCAATCACCA CTGGACAAAG  
 CCGGTGGCA TCCAGGACT AACTTAACTT TGGTGGGAAA CCTCTGAGTA TGTATCACCA CCGGTCCTCT CTGTGTCTC AGTTAGTGGT GACCGTGTTC  
 E

· S G S S I G K A F T T T L K G A Q R L A A L G E T A W D F G S V G G ·  
 3601 TCTGGAAGCA GCATTTGGCA AGCCTTTTACA ACCACCTTCA AAGGAGCGCA GAGACTAGCC GCTTAGGAG ACACAGCTTG GGAGTTTGA TCAATTGGAG  
 AGACCTTCTG CTAAACCGTT TCGAAATCT TGGTGGGAGT TTCCTCGGCT CTCTGATCGG CGAGATCTC TGTGTGAAAC CCTGAACCT AGTCAACCTC  
 E

· V F T S V G K A V E Q V F G G A F R S L F G G M S W I T Q G L L G ·  
 3701 GGTGTCTAC CTCAGTGGG AAGGCTGTCC ATCAAGTGTG CGGAGGACA TTCGCTCAC TGTTCGSAGG CATSTCTGG ATACGCAAG GATTGCTGGG  
 CCCACAGTG GAGTCAACC TTCGACAGG TAGTTACAAA GCCTCTCGT AAGGCGAGTG ACAGGCTCC GTACAGGACC TATTTGGTTC CTAACGACCC  
 E

· A L L L W M G I N A R D R S I A L T F L A V G G V L L F L S V N V ·  
 3801 GCTCTCTCTG TGTGGATGG GCATCAATGC TGGTGACAGG TCCATAGCTC TCAGTITCTT CGCAGTTGGA GGAGTCTCC TCTTCTCTC CTTGACGTTG  
 CCGAGGACC AACACCTACC CGTAGTACG AGCACTGTCC AGGTATCGAG AGTGCAAAA GCGTCAACCT CCTCAAGACG AGTAGGACAG CCACTTGCAC  
 E

3901 H A D T G C A I D I S R Q E I R C G S G V F I H N D V E A W M D R Y ·  
 CACGCTGACA CTGGGTGTGC CATAGACATC AGCCGGCAAG AGCTGAGATG TGGAGTGGG GTGTTTCATAC ACAATGATGT CGAGCCTTGG ATCGACCGGT  
 GCGGACTGT GACCCACACG GTATCTGTAG TCGCCGCTTC TCGACTCTAC ACCTTCACCT CACAAGIATG TGTACTACA CCTCCGAACC TACCTGGCCA

WN (AprME)-Rabies G PIV sequence (partial)

C protein  
 -----  
 5' UTR  
 -----

1 AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTCT GTGAGGATTA ACAACAATCA ACACAGTGGC AGCTGTTTCT TACCACGAAG ATCTCGATGT  
 TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTGTTAAT TGTGTCAAGC TCGACAAAGA ATCTGCTTC TAGAGCTACA  
 C protein  
 · K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K R A M L ·

101 CTAAGAACC AGGAGGGCCC GCAAGAGCC GGGCTGTCTA TTTGCTAAA CGCGGATGC CCGCGTGTG GTCCCTGATG GACTTAAGA GGGTATGTT  
 GATTCCTTGG TCCCTCCGGG CGGTTCTGG CCGACAGAT AAACGATTTT GCGCCTTAGG GGGCCACCA CAGCACTAA CCTGAATCT CCGCATACA  
 C protein  
 · S L I D G K G P I R F V L A L L A F F R F T A I A P T H A V L D R ·

201 GAGCCTGATC GAGGCAAGG GGCATAACG ATTGTGTGTT GCTCTCTTGG CGTCTTCAG GTTCACAGCA ATTGCTCGA CCGGAGAGT GCTGGATCGA  
 CCGGACTAG CTGCGCTTC CGGTTATGC TAAACACAC CCAGAGAACC GCAAGAAGTC CAAGTGTCTG TACGAGGCT GGGCTCGTCA CGACCTAGCT  
 C protein  
 W R G V N K Q T A M K H L L S F K K E L G T L T S A I N R R S S K Q ·

301 TCGACAGGTG TGAACAACA AACAGCGATG AACACCTTC TGAGTTCAA GAAGGACTA GGGCCTTGA CCAGTGTAT CAATCGGCGG AGCTCAARGC  
 ACCTTCCAC ACTGTITGT TGTGCTAG TTTGCGAAG ACTCAAAGT CTTCCTTGT CCTGGAAT GGTACGATA GTTAGCCGCC TCGATTTCC  
 Rabies-G signal  
 -----

|                                                                       |                                                                                                                                                                                                                                    |          |
|-----------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| C protein<br>protein                                                  | partial C signal                                                                                                                                                                                                                   | Rabies-G |
| -----                                                                 |                                                                                                                                                                                                                                    |          |
| · K K R G G K T G I A V I V P Q A I I F V P L L V F P L C F G K F P · |                                                                                                                                                                                                                                    |          |
| -----                                                                 |                                                                                                                                                                                                                                    |          |
| 401                                                                   | AAAAGAAGG AGGGGGCAG ACTGATATAG CTGTGATCTT TCCTCAGGCT CTTTGTGTTG TACCTTGCT GGTATTCCG CTTTGTGTTG GTAATTTC<br>TTTCTTGC TCCCGGTTG TGACATATC GACACTAGCA AGGAGTCCGA GAACACAC ATGGGAACA CCATAAGGG GAACGAAC CATTTAAGG<br>Rabies-G protein  |          |
| · I Y T I P D K L G P W S P I D I H H L S C P N N L V V E D E G C T · |                                                                                                                                                                                                                                    |          |
| -----                                                                 |                                                                                                                                                                                                                                    |          |
| 501                                                                   | TATCTATCC ATCCCTGATA AGCTGGGCC TTGGAGTCC ATTGATATC ACCATTTGAG CTGCCAAG ACCTCTCG TCGAGATGA AGGTGCRCT<br>ATAGATAGG TACCCACTAT TCGAGCCCG AACCTCAGG TACTATAAG TGGTAAATC GACGGGTTG TGGAGCAGC AACTCTACT TCCACGTA<br>Rabies-G protein     |          |
| N L S G F S Y M E L K V G Y I S A I K M N G F T C T G V V T E A E T · |                                                                                                                                                                                                                                    |          |
| -----                                                                 |                                                                                                                                                                                                                                    |          |
| 601                                                                   | ATCTTCTCG GATTTCTTA CTGGAGTTG AAAGTGGCT ATATTTCAG CATTAAGATG AACGCTTA CTTGTACAGG AGTCGTACC GAGCCGAGA<br>TTAGAAGAC CTAARAGAT GTACCTCAC TTTCACCGA TATAAGTCC GTATTCTAC TTGCCAAT GAACATGTC TCGCACATG CTTGCTCT<br>Rabies-G protein      |          |
| · Y T N F V G Y V T I T F K R K H F R P T P D A C R A A Y N W K M A · |                                                                                                                                                                                                                                    |          |
| -----                                                                 |                                                                                                                                                                                                                                    |          |
| 701                                                                   | CATAACAAA TTTGTTGGA TCGTCACCA CCACCTCAA GAGAAAAC TTCCGCCCA CGCTGACGC TTGCGGGCC GCTTACAAC GGAAGATGGC<br>GTATATGTT AAAGCACCT ATGCAGTGT CCGAAGTCT CTTTGTGAG AAGCGGGTT GCGGACTCG AACAGCCCG CGAATGTGA CCTTCTACG<br>Rabies-G protein     |          |
| · G D P R Y E E S L H N P Y P D Y H W L R T V K T T K E S L V I I S · |                                                                                                                                                                                                                                    |          |
| -----                                                                 |                                                                                                                                                                                                                                    |          |
| 801                                                                   | AGGAGCTCT CGATATGAG ATCTCTGCA CAACCGTAT CCTGATTACC ATTGGTGGC GACAGTCAAG ACTACCACG ASAGCTGCT CATTAATCA<br>TCCTTAGGA GCTTACTTC TTAGAGAGT GTTGGCATA GACTAATGG TACCGAGCC CTGTCAATTC TGAAGTTC TCTCAGACA GATATATG<br>Rabies-G protein    |          |
| P S V A D L D P Y D R S L H S R V F P G G N C S G V A V S S T Y C S · |                                                                                                                                                                                                                                    |          |
| -----                                                                 |                                                                                                                                                                                                                                    |          |
| 901                                                                   | CCAGCGTGG CCGATCTGA TCTTATGAT AGATCCCTGC ACAGTAGGTT TTTCTGCG GGGATTTGA GCGGTGTGC AGTATCAAGT ACCTACTGCT<br>GGTTCGCCCC GCTTAGAAT AGGAATACA CTAGGGAGC TGTATCCCA AARAGGACC CCTTAACAT CCGCACACG TCATAGTCA TGGATGACA<br>Rabies-S protein |          |
| · T N H D Y T I W K P E N P R L G M S C D I F T N S R G K R A S K G · |                                                                                                                                                                                                                                    |          |

1001 CCACTAACCA CCACTACACT ATATGGATGC CTGAGAACC TOGACTCGGT ATGAGTTGCG ACATTTTTC GAACCTACGG GGCAAGCGGG CATCTAAGGG  
 GGTGATGGT GCTGATGTGA TATACCTACG GACTCTGGG AGCTGAGCCA TACTCAACGC TGTAAAAATG CTTGAGTGCC CGTTCGCCCC GTAGATTCCC  
 Rabies-G protein  
 · S E T C G F V D E R G L Y K S L K G A C K L K L C G V L G L R L M ·

1101 CTCTGAAACA TGCGGGTTTG TTGATGAGCG GGGGTGAT AAATCTCTTA AAGGCCCTG TAAGCTGAAA CTCGTGGCG TACTGGGGCT GCGCTGATG  
 CAGACTTTG ACGCCCAAC AACTACTCGC CCCCAACATA TTTAGAGAAT TTCCCGGGAC ATTGACTTT GAGACACCGC ATGACCCCGA CGCGACTAC  
 Rabies-G protein  
 · D G T W V A M Q T S N E T K W C P P G Q L V N L H D F R S D E I E H ·

1201 GACCCACAT GGTGGCTAT GCAGACAAGC AATGAAACAA AGTGGTGTCC CCTGGTCAG CTCCTTATC TGCACGACTT TAGTCTGAC GAAATCGAGC  
 CTGCCCTGA CCCACCATA CGTCTGTTC TACTTTGTT TCACACAGG GGCACAGTC GACCAATTAG ACGTGTGAA ATCCAGACTG CTTTAGCTCG  
 Rabies-G protein  
 · L V V E E L V K K R E E C L D A L E S I M T T K S V S F R R L S H ·

1301 ACCTTGATG GAGGAACCTG CTGAGAAAAC GCGAAGATG CCGGACCGA CTTGAGAGTA TTTGACAC CAAATCCGT TCCTCAGAA GACTGAGCA  
 TGGAACACA CCTCTTCAC CACTTCTTG CGCTCTCAC GGCCTGCTT GACTCTCAT AATACTGGT GTTAGGCAA AGGAGCTTT CTGACTCGGT  
 Rabies-G protein  
 · L R K L V P G F G K A Y T I F N K T L M E A D A H Y K S V R T W N ·

1401 CCTCGAAG CTGGTCCAG GGTCCCGAA GCTTATACT ATTTTCACA AGACTCTTAT GGAGGCGAT GCCCATTATA AGTCAGTAG GACTTGAAT  
 GGCCTTTC GACCACGGT CCAAGCCCT CCGAATATG TAAAGTTGT TGTGAGATA CTTCCGCTA CGGTATAT TCACTCACT CTGAACCTTA  
 Rabies-G protein  
 · E I I P S K G C L R V G G R C H P H V N G V F F N G T I L G P D G N ·

1501 GAGATAATC CTTCAAGAG ATGTCTGAGA CTCGGTGGGA GATGCCACC CCTGTCAAT GGGGTCTTC TTAACGGAAT CATCCGGGA CCTGACCGGA  
 CTTATTAA GAGGTTTC TACAGACTCT CAGCCACCT CTACGGTGG GGTACAGTA CCCCACAAGA AATTGCTTA GTAGACCTCT GACTCCCTT  
 Rabies-G protein  
 · V L I P E M Q S S L L Q Q H M E L L V S S V I P L M H P L A D P S ·

1601 ACCTGCTGAT TCCGAGATG CAATCTTCC TTCTGAGCA ACACATGAA CTTCTGGTG CTTGAGTAT ACCCTGATG CACCCACTGG CCGACCCAG  
 TGCACGACTA AGGGCTCTG GTTGAAGGG AAGACTCTG TGTGACTTT GAGGACACA GAATCACC TGGGACTAC GTGGTGGC GCGGGGGTC  
 Rabies-G protein  
 · T V F K N G D E A E D F V E V H L P D V H E R I S G V D L G I P N ·

1701 CACTGTCTC AAAAATGGG ATGAGGCCGA AGACTTTGG GAATTCACC TGCCCGATG ACAAGAAAGG ATATCTGGG TAGACCTGG CCTTCTAAT  
 GTGACACA TTTTACCCT TACTCCGCT CTGAAACAC CTTCAAGTGG AGGGCTACA TGTGCTTCC TATACACCT ATCTGGACC GAAAGGATA  
 Rabies-G protein  
 · W G K Y V L L S A G A L T A L M L I I F L M T C W R R V N R S E P T ·

1801 TGGGTAAAT ACGTGTCTT GAGTCCGGT GCCTTGACCG CTTGATGCT GATCATTTT CTGATGACT GCTGGCGAG GGTGAATCC TCCGAGCCGA  
 ACCCATTA TGCAGAGGA CTCACGCCA CCGAAGTGG GAAACTACGA CTAGTAAAA GACTACTGGA CACCGGCTC GACTTACGG AGGCTCGGCT  
 Rabies-G protein  
 · Q E N L R G T S R E V S V T P Q S G K I I S S W E S Y K S G G E T ·

1901 CACAGCACA TCTCAGAGG ACAGGCCGG AAGTAAAGT GACTCCGAA TCTGGCAAGA TTATTAGTAG TGGGAGGT TACAAGTCTG GAGGAGAGAC  
 GTGCTGTT AGACTTCCC TGTCCGGCC TCAATCACA CTGACCCCT AGACCGTCT AATAATCAT AACCTCTCA ATGTTGACG CTTCTCTCTG  
 FMDV 2A NS1 signal  
 Rabies-G protein preNS1 signal  
 · C L N F D L L K L A G D V E S N P C P A R D R S I A L T F L A V G ·

2001 TGGTTCAAT TTTGATCTC TCAAACTTC AGGCGATGA GAATCAAATC C:GGACCGC CCGGACAGG TCCATAGCTC TCACGTTCT CCGAGTTGGA  
 ACCCAACTA AAAGTAGAG ACTTTGAAC TCCGCTACAT CTTAGTTAG GACCTGGCG GGCCTCTCC AGGTATCGAG AGTGCAAGA GCTCAACT  
 NS1  
 NS1 signal  
 · G V L L F L S V N V H A D T G C A I D I S R Q E L R C G S G V F I H ·

2101 GGAGTCTGC TCTTCTCTC CGTGAACGC CACGCTGACA CTGGGTGTC CATAGACATC AGCCGCAAG AGCTGAGTG TCGAAGTGA GTGTTTATC  
 CCTCAAGAG AGAAGGAGAG GCACTTGAC CTCCACTGT GACCCACAG GTATCTGTAG TGGCGCTC TCGACTAC ACCTCACCT CACAAGTATG

Sequence Appendix 4 (continued)

PIV-WNV helper ΔNS1

C

5' UTR

-----

1 AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTTT GTGAGGATTA ACACAATTA ACACAGTCCG AGCTGTITCT TAGCACGAAG ATCTCGATGT M S  
TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTGTTAAT TGTGTCACGC TCGACAAAGA ATCGTGTCTC TAGAGCTACA  
C

-----

101 · K K P G G P G K S R A V N M L K R G M P R V L S L I G L K R A M L ·  
CTAAGAAACC AGGAGGGCCC GGCAAGAGCC GGGCTGTCAA TATGCTAAA CGCGGAATGC CCCGCGTGT GTCCTTGATT GGACTTAAGA GGGCTATGTT ·  
GATTCITTTG ICTCCCGGG CCGTCTCGG CCCGACAGTT ATACGATTTT GCGCCTTACG GGGCGCACAA CAGGAAGTAA CCTGAATTTT CCCGATACAA  
C

-----

201 · S L I D G K G P I R F V L A L L A F F R F T A I A P T R A V L D R ·  
GAGCCTSATC GACGSCAAGG GGCCAATAGC AITTTGTGTT GCICCTTTGG CGTTCCTCAG GTTCACAGCA ATTGCTCCGA CCCGAGCAGT GCTGGATCGA  
CTCGGACTAG CTGCGTTC CCGGTTATGC TAAACACAAC CGAGAGAACC GCAAGAATC CAAGTGTCTT TAACGAGGCT GGGCTCGTCA CGACCTAGCT  
C

-----

301 W R G V N K Q T A M K H L L S F K K E L G T L T S A I N R R S S K Q ·  
TGAGAGGTTG TGAACAAACA AACAGCGATG AAACACCTTC TGAGTTTCAA GAAGGAAGTAA GGGACCTTGA CCAGTGTCTAT CAATCGGCGG AGCTCAAAC  
ACCTCTCCAC ACTTGTITGT TTGTGCTAC TTGTGGAAG ACTCAAAGTT CTTCCTTGAT CCCTGGAAC TGTACAGATA GTTAGCCGCC TCGAGTTTTG  
Signal peptide

-----

C prM

-----

401 · K K R G G K T G I A V M I G L I A S V G A V T L S N F Q G K V M M ·  
AAAAGAAAAG AGGAGGAAG ACCCGAATTG CAGTCATGAT TGGCCTGATC GCCAGCGTAG GAGCAGTTAC CCTCTTAC TTCCAAGGA AGTGATGAT  
TTTTCTTTTC TCCTCCTTC TGGCCTTAC GTCAGTACTA ACCGACTAG CCGTCGCATC CTCGTCAATG GGAGAGATTG AAGTTCCCT TCCACTACTA  
prM

-----

501 · T V N A T D V T D V I T I P T A A G K N L C I V R A M D V G Y M C ·  
GACGGTAAAT GCTACTGACG TCACAGATGT CATCACGATT CCAACAGCTG CTGGAAGAA CCTATGCATT GTCAGAGCAA TGGATGTGGG ATACATGTGC  
CTGCCATTTA CGATGACTGC AGTGTCTACA GTAGTCTAA GGTGTGCGAC GACCTTICTT GGATACGTA CAGTCTGTT ACCTACACC TATGTACACG  
prM

-----

601 D D T I T Y E C F V L S A G N D P E D I D C W C T K S A V Y V R Y G ·  
GATGAACTA TCACCTTATGA ATGCCAGTG CTGTCCGCTG GTAATCATCC AGAAGACATC GACTGTGGT GCACAAAGTC AGCAGTCTAC GTCAGGTATG  
CTACTATGAT AGTGAATACT TACGGGTCAC GACAGCCGAC CATTACTAGG TCTTCTGTAG CTGACAACCA CGTGTTCAG TCCTCAGATG CAGTCCATAC  
prM

-----

701 · R C T K T R H S R R S R R S L T V Q T H G E S T L A N K K G A W M ·  
GAAGATGCAC CAAGACACG CACTCAAGAC GCAGTCGGAG GTCACIGACA GTGCAGACAC ACGGAGAAG CACTTAGCG AACAAGAAGG GGGCTTGGAT  
CTTCTAAGTG GTTCTGTGGG GTGAGTCTG CGTCAGCCTC CAGTACTGT CACGCTCTGT TCCTCTTTC GTGAGATCGC TTGTTCTTCC CCGCACCTA

p1M

801 · D S T K A T R Y L V K T E S W I L R N P G Y A L V A A V I G W M L  
GGACAGCACC AAGCCACAAA GGTATTGGT AAAAACAGAA TCATGGATCT TGAGGAACCC TGGATATCC CTGGTGGCAG CCGTCATTGG TTGGATGCTT  
CCTGTCGTGG TTCCGGTGT CCATRAACCA TTTTGTCTT AGTACCTAGA ACTCCTTGGG ACCIATACGG GACCACCCTC GGCAGTAACC AACCTACGAA  
E

p2M

901 · G S N T M Q R V V F V V L L L L V A P A Y S F N C L G M S N R D F L ·  
GGGAGCAACA CCATGCAGAG AGTTGTGTTT GTCCGTGCTAT TGCTTTTGGT GGCCCGAGCT TACAGCTTTA ACTGCCCTGG AATGAGCAAC AGAGACTTCT  
CCCTCGTTGT GGTACGTCCT TCAAGACAAA CAGCAGGATA ACGAAABACCA CCGGGGTGCA ATGTCGAAT TGACGGAAAC TTACTCGTTG TCTCTGAAGA  
E

1001 · E G V S G A T W V D L V L E G D S C V T I M S K D K P T I D V K M ·  
TGGAGGAGT CTC?GGAGCA ACATGGGTGG ATTTGGTCTT CGAAGGGCAG AGCTCCGTGA CTATCATGTC TAAGGACAAG CCTACCATCG ATGTGAAGAT  
ACCTTCCTCA CAGACCTCGT TGTACCCACC TAAACCAAGA GCTCCGCTG TCGAGCCTACT GATAGTACAG ATTCCIGTTC GGAIGGTAGC TACACTTCTA  
E

1101 · M N M E A A N L A E V R S Y C Y L A T V S D L S T K A A C P A M G ·  
GATGAATATG GAGGCGGCA ACCTGGCAGA GTCCCGCAGT TATTGCTATT TGGCTACCGT CAGCGATCTC TCCACCAAAG CTCGCTGCC GGCATGGGA  
CTACTTATAC CTCGCGCGT TGGACCGTCT CCAGGCCCTCA ATAACGATAA ACGGATGCA GTCCGTAGAG AGTGGTTC GAGCAGCGG CCGTACCT  
E

1201 · E A H N D K R A D P A F V C R Q G V V D R S W G N G C G L F G K G S ·  
GAGCTCACA ATGACAAACG TGCTGACCCA SCTTTTGTGT CGACACAAG AGTGGTGGAC AGGGGCTGG GCAACGGCTG CCGACTATT GGCAAGGAA  
CTCGAGTGT TACTGTTTGC ACGACTGGT CGAAACACA CGTCTGTTC TCACCACCTG TCCCAGACC CGTTGCCGAC GCGTGATAAA CCGTTCTCTT  
E

1301 · I D T C A K F A C S T K A I G R T I L K E N I K Y E V A I F V H G ·  
GCATTGACAC ATGCGCCAAA TTTCCCTGCT CTACCAAGGC AATAGGAAGA ACCATTTGA AAGAGAATAT CAAGTACGAA GIGGCCATT TTGTCCATGG  
CGTAACCTG TACCGGTTTT AAACGGACGA GATGGTTCCG TTATCCTTCT TGGTAAACT TTCTTTATA GTTCATGCTT CACCGTAAA AACAGGTACC  
E

1401 · P T T V E S H G N Y S T Q V G A T Q A G R F S I T P A A P S Y T L ·  
ACCAACTACT GTGGAGTCGC ACGGAACTA CTCCACACAG GTTGGAGCCA CTGAGGCAGG GAGATTGAGC ATCACTCCTG CCGCCCTTC ATACACACTA  
TGGTTGATGA CACCTCAGCG TGCCTTTGT GAGGTGTCT CAACTCCGCT GAGTCCGTCC CTCTAAGTGC TAGTGGAGAC GCCCGGGAAG TATGTGTGAT  
E

1501 · K L G E Y G E V T V D C E P R S G I D T N A Y Y V M T V G T K T F L ·  
AAGCTTGGAG AATATGGAGA GGTGACAGTG GACTGTGAAC CACGGTCAGG GATTGACACC AATGCATACT ACGTGTGAC TGTGGAACA AAGACGTTCT  
TTCGAACCTC TTATACCTCT CCACGTCTAC CTGACACTTG GTGCCAGTCC CTAAC?TGG TTACGTATGA TGCCTACTG ACAACCTTGT TTCTGCAAGA  
E

1601 · V H R E W F M D L N L P W S S A G S T V W R N R E T L M E F E E P ·  
TGGTCCATCG TGAGTGGTTC ATGGACCTCA ACCTCCCTTG CAGCACTGCT GGAAGTACTG TGTGGAGGAA CAGAGAGACG TTAATGGAGT TTGAGGAACC  
ACCAGGTAGC ACTCACCAAG TACCTGGAGT TGGAGGGAAC CTCGTACGCA CCI?CATGAC ACACCTCCTI GTCTCTCTGC AATTACCTCA AACTCCTTGG  
E

1701 · H A T K Q S V I A L G S Q E G A L H Q A L A G A I P V E F S S N T ·  
ACAGCCACG AAGCAGTCTG TGATAGCATT GGGCTACAA GAGGAGCTC TGCATCAAGC TTTGGCTGGA GCCATTCCTG TGGAAATTTT AAGCAACACT  
TGTCCGGTGC TTCTCAGAC ACTATCTGTA CCCGAGTGT CTCCCTCGAG ACGTAGTTCG AAACCGACCT CCGTAAGGAC ACCTTAAAG TCTGTGTGA  
E

```

V K L T S G H L K C R V K M E K L Q L K G T T Y G V C S K A F K F L
1801 GTC AAGTGA CGTCGGCTCA TTG AAGTGT AGAGTGAAGA TGGAAAAAT CCAGTTGAAG GGAACAACCT ATGGCGTCTG TTC A AAGGCT TTC AAGTTC
CAGTCAACT GCAGCCAGT AAACCTCACA TCTCACTTCT ACCTTTTAA CTCAACTTC CTTTGTGGA TACCGCAGAC AAGTTTCCGA AAGTTC A AAG
E
-----
G T P A D T G H G T V V L E L Q Y T G T D G P C K V P I S S V A S
1901 TTGGGACTCC CGCAGACACA GGTCCAGGCA CTGTGGTGT GGAATTCAG TACACTGGCA CGGATGGACC TTGCAAAGTT CCTATCTCGT CAGTGGCTTC
AACCTGAGG GGTCTGTGT CCAGTCCCT GACACCACAA CCTTAACGTC ATGTGACCGT GCCTACCTGG AACGTTCFAA GGATAGAGCA GTCACCCGAG
E
-----
L N D L T P V G R L V T V N P F V S V A T A N A K V L I E L E P P
2001 ATTGAACGAC CTAACGCCAG TGGGCAGATT GGTCACTGTC AACCCITTTG TTTCACTGGC CACGGCCAAAC GCTAAGTCC TGATTGAATT GGAACCAACC
TAACTTGCTG GATTGCGGTC ACCCGTCTAA CCAGTGACAG TTGGGAAAC AAAGTCACCG GTCCCGGTTG CGATTCCAGG ACTAACTTAA CCTTGGTGGG
E
-----
F G D S Y I V V G R G E Q Q I N H H W H K S G S S I G K A F T T T L
2101 TTGGGACT CATACTAGT GGTGGGACAGA GGAGAACAC AGATCAATCA CCCTGGCAC AAGCTGGAA GCACCATTTG CAAGCCCTT ACAACCAACC
AAACCTGTA GTATGTATCA CCACCCGCT CCTCTTGTG TCTAGTACT GGTGACCGTG TTCAGACTT CTCTGAACC GTTTCGAAA TGTGGTGGG
E
-----
K G A Q R L A A L G D T A N D F G S V G G V F T S V G K A V H Q V
2201 TCAAGSASC SCAGAGACTA GCCCTCTAG GAGACACAGC TTGGGACTTT GSATCAGTGG GAGGGGTGT CACCTCAGTT GGGAGGGCTG TCCATCAAGT
AGTTCCTCG CGTCTCTGAT CGCCGAGATC CTCTGTGTG AACCCGTAAA CTTAGTCAAC CTCGCCACAA GTGGAGTCAA CCTTCCGAC AGGTAGTCA
E
-----
F G S A F R S L F G G M S W I T Q G L L G A L L L W M G I N A R D
2301 GTTCGGAGGA GCATTCGGT CACTGTCGG AGGCATGTC TGGATAACGC AAGGATTGCT GGGGGCTCTC CTGTTGTGGA TGGGCATCAA TGCTCGIGAC
CAAGCCTCT CGTAAGCGA GTGACAAGCC TCCGTACAGG ACCTATTGCG TTCTTAACGA CCCCCGAGAG GACAAACACT ACCCGTAGTT ACGAGCACTG
deleted NS1
E
-----
R S I A L T F L A V G G V L L F L S V N V H A D T G I H R G P A T R
2401 AGTCCATAG CTCTCAGTT TCTCGCAGTT GGAGGAGTTC TGCTCTTCTT CTCCGTGAAC GTGCACGCTG ACCTGGGAT CCACCGTGA CCTGCCACTC
TCCAGGTATC GAGAGTGCAA AGAGCGTCAA CCTCCTCAAG ACGAGAGGA GAGGCACTTG CACGTGGCAG TGTGACCTTA GGTGGCACTT GGACGGTGC
deleted NS1
-----
T T T E S G K L I T D W C C R S C T L P P L R Y Q T D S G C W Y G
2501 GCACCCACAG AGAGAGCGGA AAGTTGATAA CAGATTGGTG CTCCAGGAGC TGCACCTTAC CACCACTCG CTACCAACT GACAGCGGCT GTTGGTATGG
CCTGCTGCTG TCTCTGCGCT TTCAACTAT CTCAACCCAC GACCTCTTCC ACSTGGATG GTGGTGCAGC GATGGTTGA CTGTCCCGGA CAACCATACC
deleted NS1
-----
NS2A
-----
M E I R P Q R H D E K T L V Q S Q V N A Y N A D M I D P F Q L G L
2601 TATGGAGATC AGACCACAGA GACATGATGA ANAGACCCTC GTGACGTCAC AAGTGAATGC TTATAATGCT GATATGATTG ACCCTTTTCA GTTGGGCCCTT
ATACCTCTAG TCTGGTGTCT CTGACTACT TTCTGGGAG CACGTCAGTG TTCACTTACG AATATTACGA CTATACTAAC TGGGAAAGT CAACCCGGA

```

Sequence Appendix 5

PIV-WNV( $\Delta$ prME)/RSV-F

C protein  
~~~~~

5' UTR

1 AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTIT GTGAGGATTA ACAACAATTA ACACAGTCCG AGCTGTTTCT TAGCACGAAG ATCTCGATGT M S
TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTGTGTAAT TGTGTCAAGC TCGACAAAGA ATCGTGCTTC TAGAGCTACA

C protein

101 . K K P G C P G K S R A V Y L L K R G M P R V L S L I G L K R A M L .
CTAAGAAACC AGGAGGGCCC GGCAGAGCC GGGCTGTCTA TTGCTAAAA CGCGGAATGC CCCGCGTGTG GTCTTGATT GGACTTAAGA GGGCTATGTT
GATTCCTTGG TCCTCCCGGG CCGTTCCTGG CCCGACAGAT AAACGATTTT GCGCCTTACG GGGCGCACAA CAGGAACTAA CCTGAATTCT CCCGATACAA

C protein

201 . S L I D G K G P I R K V L A L L A F F R F T A I A P T R A V L D R
GAGCCTGATC GACGGCAAGG GGCCAATACG AITTTGTTTG GCCTCTTTGG CGTTCCTTACG STTCACAGCA ATTGCTCCGA CCCGAGCAGT GCTGGATCGA
CTCGGACTAG CTGCGGTTC CCGGTTATGC TAAACACAAC CGAGAGAACC GCAAGAAGTC CAAGTGTCTT TAACGAGGCT GGGCTCGTA CGACCTAGCT

NS3 cleavage
~~~~~

C protein

301 W R G V N K Q T A M K H L L S F K K E I G T L T S A I N R R S S K Q .  
TGGAGAGGTG TGAACAACA AACAGCGATG AAACACCTTC TGAGTTTCAA GAACGAACATA GGGACCTTGA CCAGTGTAT CAATCGGCGG AGCTCAAGCG  
ACCTCTCCAC ACTTGTTTGT TTGTGCTAC TTTGTGGAAG ACICAAAGTT CTTCCTTGAT CCTGGAACG GGTCCAGATA GTTAGCCGCG TCGAGTTTGC

F signal

NS3 cleavage

F1

401 . K K R G G E L L I L K A N A I T T I L T A V T F C F A S G Q N I T .  
AAAAGAAGCG AGGGGGCGAG TTGCTAATCC TCAAAGCAAA TGCAATTACC ACAATCCTCA CTGCAGTCAC ATTTTGTITT GCTTCTGGTC AAAACATCAC  
TTTTCTTCGC TCCCGCGCTC AACGATTAGG ASTTTCGTTT ACGTTAATGG TGTTAGGAGT GACGTCAGTG TAAACAAAA CGAAGACCAG TTTTGTAGTG

F1

501 . E E F Y Q S T C S A V S K G Y L S A L R T G W Y T S V I T I E L S .  
TGAAGAAATT TATCAATCAA CATGCAGTGC AGTTAGCAAA GGCTATCTTA GTGCTCTGAG AACTGGTTGG TATACCAGTG TTATAACTAT AGAATTAAGT  
ACTTCTTAAA ATAGTTAGT STACGTCACG TCAATCGTTT CCGATAGAA CAGGAGACTC TTGACCAACC ATATGGTCAC AATATGATA TCTTAATTCA

F1

601 N I K E N K C N G T D A K V K L I K Q E L D K Y K N A V T E L Q L L .  
AATATCAAGG AAAATAAGTG TAATGGAACA GATGCTAAGG TAAATTTGAT AAAACAAGAA TTAGATAAAT ATAAAAATGC TGTRACAGAA TTGCAGTTGC  
TTATAGTTCC TTTTATTCAC ATTTACTTGT CTACGATTC APTTAACTA TTTTGTCTT AATCTATTTA TATTTTACG ACATGTCCT AACGTCACG

F1

701 . M Q S T P P T N N R A R R E L P R F M N Y T L N N A K K T N V T L .  
TCATGCAAG CACACCACA ACAACAATC GAGCCAGAAG AGAATACCA AGGTTTATGA ATTATACACT CACAATGCC AAAAACAACA ATGTACATT  
AGTACGTTT GTGTGGTGGT TGTTTGTAG CTCGCTCTC TCTTGTGGT TCCAAATACT TAATACTGTA GTTGTACCG TTTTGTGGT TACATTGTAA

F1

-----

```

      · S K K R K R R F L G F L L G V G S A I A S G V A V S K V L H L E G
801 AAGCAGAAA AGAAAAAGAA GATTTCTTGG TTTTTTGTTA GGTGTTGGAT CTGCAATCGC CAGTGGCGTT GCTGTATCTA AGGTCCCTGCA CCTAGAAGGG
      TTCGTTCTTT TCCTTTCTT CTAAGAAGCC AAAAAACAAT CCACAACCTA GACGTTAGCG GTCACCGCAA CGACATAGAT TCCAGGACGT GGATCTTCCC
      F2
      · E V N K I K S A L L S T N K A V V S L S N G V S V L T S K V L D L K
901 GAAGTGAACA AGATCAAAAG TGCTCTACTA TCCACAACA AGGCTGTAGT CAGCTTATCA AATGGAGTTA GTGTCCTAAC CAGCAAAGTG TTAGACCTCA
      CTTCACTTGT TCTAGTTTTC ACGAGATGAT AGGTTTGT TCCGACATCA GTCGAATAGT TTACCTCAAT CACAGAATFG GTCGTTTCAC AATCTGGAGT
      F2
      · N Y I D K Q L L P I V N K Q S C S I S N I E T V I E F O Q K N N R
1001 AAAACTATAT AGATAACAA TTGTACCTA TTGTGAACAA GCAAGCTGC AGCATATCAA ATATAGAAC TGTGATAGAG TTCCAAACAA AGACAACAG
      TTTTGATATA TCTATTGTT AACAATGGAT AACACTTGT CTTTTGACG TCGTATAGT TATATCTTTG ACACIATCTC AAGGTTGTTT TCTTGTGTG
      F2
      · L L E I T R E F S V N A G V T T P V S T Y M L T N S E L L S L I N
1101 ACTACTAGAG ATTACCAGG AATTTAGTGT TAATGCAGGT GTAACACAC CTGTAAGCAC TTACATGTTA ACTAATAGTG AATTATTCTC ATTAATCAAT
      TGATGATCTC TAATGGTCCC TTAATCACA ATTACGTCCTA CATTGATGTG GACATTCTGT AATGTACAAT TGATTATCAC TTAATAACAG TTAATTAGTTA
      F2
      · D M P I T N D Q K K L M S N N V Q I V R Q Q S Y S I M S I I K E E V
1201 GATATGCCIA TAACAAATCA TCAGAAAAAG TTAATGTCCA ACAATGTTCA AATAGTTAGA CAGCAAAGTT ACTCTATCAT GTCCATAATA AAAGAGGAAG
      CTATACGGAT ATTGTTTACT AGTCTTTTC AATTACAGGT TGTACAAGT TTATCAATCT GTCGTTTCAA TGAGATAGTA CAGGTATTAT TTTCTCTTC
      F2
      · L A Y V V Q L P L Y G V I D T P C W K L H T S P L C T T N T K E G
1301 TCTTAGCATA TGTAGTACAA TTACCACTAT ATGGTGTIAT AGATAGACCC TGTGGAAAC TACACACATC CCTCTATGT ABRACCAACA CAAAGAGS
      AGAATCGIAT ACATCATGTT AATGGTGATA TACCACAATA TCTATGTGGG ACAACCTTG ATGTGTSTAG GGGAGATACA TGTGGTGTG TTTTCTTCC
      F2
      · S N I C L T R T D R S W Y C D N A G S V S F F P Q A E T C K V Q S
1401 GTCACACATC TGTTTAACAA GAACGACAG AGGATGGIAC TGTGACAATG CAGGATCAGT ATCTTCTCTC CCACAAGCTG AACACATGTA AGTTCAATCA
      CAGGTTCTAG ACAAATTGTT CTTGACTGTC TCTACCATG ACACGTGTAC GTCCTAGTCA TAGAAAGAAG GGTGTTCCAC TTTGTACATT TCAAGTTAGT
      F2
      · N R V F C D T M N S L T L P S E I N L C N V D I F N P K Y D C K I M
1501 AATCGAGTAT TTTGTGACAC AATGAACAGT TTAACATTAC CAAGTGAAAT AAATCTCTGC AATGTGACA TATTCAACCC CAATATGAT TGTAATAATA
      TTAGCTCATA AAACACTGTG TTACTTGICA AATTGTAATG GTTCACTTIA TTTAGAGACG TTACAACCTG ATAAGTTGGG GTTTTACTA ACATTTTAT
      F2
      · T S K T D V S S S V I T S L G A I V S C Y G K T K C T A S N K N R
1601 TGACTTCAA AACAGATGTA AGCAGTCCG TTATCACATC TCTAGGAGCC ATTGTGTCTT GCTATGGCAA AACTAATGT ACAGCATCCA ATAAATATC
      ACTGAAGTTT TTGTCTACAT TCGTCCGAGC AATAGTGTAG AGATCCCTCG TAACACAGTA CGATACCGTT TTGATTTACA TGCTGAGST TATTTTTAGC
      F2
      · G I I K T F S N G C D Y V S N K G M D T V S V G N T L Y Y V N K Q
1701 TGAATCATA AAGACATTTT CTACCGGGTG CGATTATGTA TCAATAAAG GATGGACAC TGTGCTGTG GGTAAACACAT TATATTATGT AAATAGGCAA
      ACCTTAGTAT TTCTGTAATA GATTGCCAC GCTAATACAT AGTTTATTTT CCTACCTGTG ACACAGACAT CCATTGTGTA ATATAATACA TTTATTCGTT
      F2
      · E G K S L Y V K G E P I I N F Y D P L V F P S D E F D A S I S Q V N
1801 GAAGGTAAA GTCTCTATGT AAAAGGTGAA CCAATAATA ATTTCTATGA CCATTAGTA TTCCCTCTG ATGAATTTGA TGCAI'CAATA TCTCAAGTCA

```

CITCCATTTT CAGAGATACA TTTTCCACTT GGTATTATT TAAAGATACT GGGTAATCAT AAGGGGAGAC TACTTAAACT ACGTAGTTAT AGAGTTCAGT  
F2

TM Domain  
E K I N Q S L A F I R K S D E L L H N V N A G K S T T N I M I T T  
1901 ACGAGAAGAT TAACCAAGAGC CTAGCATTTA TTCGTAATC CGATGAATTA TTACATAATG TAAATGCTGG TAAATCCACC ACAAATATCA TGATAACTAC  
TGCTCTTCTA ATTGGTCTCG GATCGTAAAT AAGCAITTAG GCTACTTAAT AATGTATTAC ATTTACGACC ATTTAGGTGG TGTTTATAGT ACTATTGATG  
TM Domain

Cytoplasmic Tail  
I I I V I I V I L L S L I A V G L L L Y C K A R S T P V T L S K D  
2001 TATAATTATA GTGATTATAG TAATATTGTT ATCATTAATT GCTGTGGAG TGCTCTTATA CTGTAAGGCC AGAAGCACAC CAGTCACTCT AAGCAAAGAT  
ATATTAAAT CACTAATATC ATTATAACA TAGTAATTAA CGACACCTG ACGAGACTAT GACTTCCGG TCTTCGTGTG GTCAGTCTGA TTCGTTCTTA  
FMDV 2A

domain of WNV E (split) Transmembrane  
Cyttoplasmic Tail pre E/NS1 signal

Q L S G I N N I A F S N N F D L L K L A G D V E S N P G P A R D R S  
2101 CAACTGAGTG GTATAAATA TATTGCATT AGTAACAATT TTGATCTGCT CAACCTGCA GCGGATGTAG AATCAAATCC TGGACCCGCC CGGGACAGGT  
GTTGACTCAC CATATTATT ATAACSTAAA TCATTGTAA AACTAGACGA GTTTGAACGT CGCTACATC TTAGTTTAGG ACCTGGGCGG GCCTGTCCA  
NS1

Transmembrane domain of WNV E (split)  
I A L T F L A V G G V L L F L S V N V H A D T G C A I D I S R Q  
2201 CCATAGCTCT CACGTTTCTC GCAGTTGGAG GACTTCTGCT CTTCCTTCC GTGAACSTGC ACGCTGACAC TGGGTGTGCC ATAGACATCA GCCGCCAA  
GGTATCGAGA GTGCAAAGAG CGTCAACCTC CTCAGACGA GAAGGAGAGS CACTTSCAG TCCGACTGTG ACCCACACGG TATCTGTAGT CGGCCGTT

PIV-WNV(ΔCprME)/RSV-F

deleted C protein  
5' UTR  
M S  
1 AGTAGTTCGC CTGTGTGAGC TGACAAACTT AATAGTGTTC GTGAGGATTA ACAACAAATTA ACACAGTGGC AGCTGTTTCT TAGCACGAAG ATCTCGATGT  
TCATCAAGCG GACACACTCG ACTGTTTGAAT TCATCACAAA CACTCCTAAT TGTGTGTAAT TGTGTCACGC TCGACAAAGA ATCGTGTCTC TAGAGCTACA  
deleted C protein  
NS3 cleavage  
K K P G G P G K S R A V N M L K R G M P R V L S L I G L K Q K K R

101 CTAAGAAACC AGGAGGGCCC GGCAAGAGCC GGGCTGTCAA TATGCTAAAA CGCGGAATGC CCGCGGTGTT GTCCTTGATT GGACTTAAAC AAAAGAAAGCC  
 GATTCTTTGG TCCTCCCGGG CCGTTCCTCG CCGACAGTT ATACGATTTT CCGCCTTACG GGGCGCACAA CAGGAACTAA CTTGAATTCG TTTTCTTCGG  
 F signal

-----  
 NS3 cleavage F1  
 ~~~~~

201 · G G E L L I L K A N A I T T I L T A V T F C F A S G Q N I T E E F
 AGGGGGCGAG TTGCTAATCC TCAAGCAAAA TCGAATTACC ACAATCCTCA CTGCAGTCAC ATTTTGTITT GCTTCTGGTC AAAACATCAC TGAAGAATTT
 TCCCCCGCTC AACGATTAGG AGTTTCGTTT ACGTTAATGG TGTAGGAGT GACGTCAGTG TAAACAAAA CGAAGACCAG TTTTGTACTG ACTTCTTAAA
 F1

 301 · Y Q S T C S A V S K G Y L S A L R T G W Y T S V I T I E L S N I K E ·
 TATCAATCAA CATGCACTGC AGTTAGCAAA GGCTATCTTA GTGCTGTGAG AACCTGGTGG TATACCAGTG TTATAACTAT AGAATTAAGT AATATCAAGG
 ATAGTTAGTT GTACGTCACG TCAATCGTTT CCGATAGAAT CACGAGACTC TTGACCAACC ATATGGTCAC AATATTGATA TCTTAATCA TTATAGTTCC
 F1

 401 · N K C N G T D A K V K L I K Q E L D K Y K N A V T E L Q L L M Q S ·
 AAAATAAGTG TAATGCAACA GATGCTAAGG TAARATTGAT AAAACAAGAA TTAGATAAAT ATAAAAATGC TGTAAACAGAA TTGCAGTTGC TCATGCAAGG
 TTTTATTAC ATTACCTTGT CTACGATTC ATTTAACTA TTTTGTCTT AATCTATTTA TATTTTACG ACATTTGCTT ACGTCAACG AGTACGTTTC
 F1

 501 · T P P T N N R A R R E L P R F M N Y T L N N A K K T N V T L S K K ·
 CACACCACCA ACAAACAATC GAGCCAGAAG AGAATACCA AGGTTTATGA ATTATACACT CAACAATGCC AAAAAACCA ATGTAACATT AAGCAAGAAA
 GTGTGGTGGT GTTGTGTAG CTCGGTCTTC TCTTGATGGT TCCAATACT TAATATGIGA GTTGTACGG TTTTGTGGT TACATTGTA TCCGTTCTTT
 F1

 F2

 601 R K R R F L G F L L G V G S A I A S G V A V S K V L H L E G E V N K ·
 AGGAAAAGAA GATTTCCTGG TTTTGTGTA GGTGTGGAT CTGCAATCCG CAGTGGCGTT GCTGTATCTA AGTCTCTGCA CTTAGAAGGG GAAGTGAACA
 TCCTTTTCTT CTAAGAACC AAAAACAAAT CCACAACCTA GACGTTAGCG GTCACCGCAA CGACATAGAT TCCAGGAGCT GGATCTTCCC CTTCACTTGT
 F2

 701 · I K S A L L S T N K A V V S L S N G V S V L T S K V L D L K N Y I ·
 AGATCAAAAG TGCTCTACTA TCCACAACA AGGCTCTACT CAGCTTATCA AATGGAGTTA GTGTCTTAA CAGCAAAAGTG TTAGACCTCA AAAACTATAT
 TCTAGTTTC ACGAGATGAT AGGTGTTTGT TCCGACATCA CTCGAATAGT TTACCCTAAT CACAGAAITG GTCGTTTAC AATCTGGAGT TTTTATATA
 F2

 801 · D K Q L L P I V N K Q S C S I S N I E T V I E F Q Q K N N R L L E ·
 AGATAACAA TTGTACTCTA TTGTGAACAA GCAAGCTGC AGCATATCAA ATATAGAAAC TGTGATAGAG TTCCAACAAA AGAACACAG ACTACTAGAG
 TCTATTTGTT AACAATGGAT AACACTTGT CGTTTCGAGC TCGTATAGTT TATATCTTGG ACACATATCT AAGGTTGTTT TCTTGTGTC TGAIGATCTC
 F2

 901 I T R E F S V N A G V T T P V S T Y M L T N S E L L S L I N D M P I ·
 ATTACCAGGG AATTAGTGT TAATGCAGGT CTAACACAC CTGTAAGCAC TTACATGTTA ACTAATAGTG AATTATTGTC ATTAATCAAT GATATGCCTA
 TAATGGTCCC TTAATACACA ATTACCTCCA CATTCATGTG GACATTCGIG AATGTACAAT TGATTATCAC TTAATAACAG TAATTAGTTA CTATACCGAT
 F2

 1001 · T N D Q K K L M S N N V Q I V R Q Q S Y S I M S I I K E E V L A Y ·
 TAACAAATGA TCAGAAAAG TTAATGTCCA ACAATCTCA AATAGTTAGA CAGCAAGIT ACTCTATCAT GTCCATAATA AAAGAGGAAG TCTTAGCATA
 ATTTGTTACT AGTCTTTTTT AATTACAGGT TGTTACAAGT TTATCAATCT GTCGTTTCAA TGAGATAGTA CAGGTATTAT TTTCTCTTC AGAATCGTAT
 F2

 · V V Q L P L Y G V I D T P C W K L H T S P L C T T N T K E G S N I

```

1101 TGTTAGTACAA TTACCACATAT ATGGTGTAT ATAGTACACCC TGTTGGAAAC TACACACATC CCCTCTATGT ACAACCAACA CAAAGAGAGG GTCCACATC
ACATCATGTT AATGGTGATA TACCACAATA TCATATGGGG ACAACCTTTG ATGTGTGTAG GGGAGATACA TGTGGTGTGT GTTTTCITCC CAGGTTGTAG
F2
C L T R T D R G W Y C D N A G S V S F F P Q A E T C K V Q S N R V F
1201 TGTTTAACAA GNACTGACAG AGGATGGTAC TGTCACAATG CAGGATCAGT ATCTTTCTTC CCACAAGCTG AACATGTAA AGTTCAATCA AATCGAGTAT
ACAAATTTGT CTGACTGTG TCCTACCATG ACACGTGTTAC GTCCTAGTCA TAGAAAGAAG GGTGTTCGAC TTTGTACATT TCAAGTTAGT TTAGTTCATA
F2
C D T M N S L T L F S E I N L C N V C I F N P K Y D C K I M T S K
1301 TTTGTGACAC AATGAACAGT TTAACATTAC CAAGTGAAT AATCTCTGCG AATGTTGACA TATTCACCCC CAATATGAT TGTAATAATG TGACTTCAAA
AAACACTGTG TTACTTGICA AATGTGAATG GTTCACITTA TTTAGAGAGC TTACAACCTGT ATAAGTTGGG GTTTATACTA ACATTTTAAAT ACTGAAGTTT
F2
T D V S S S V I T S L G A I V S C Y G K T K C T A S N K N R G I I
1401 AACAGATGTA AGCAGCTCG TTATCACATC TCTAGGAGCC AATGTTGTCAT GCTATGGCAA AACTAAATGT ACAGCATCCA ATAAAAATCG TGGAAATCATA
TTGTCTACAT TCGTCGAGGC AATAGTGTAG AGATCCCTGG TAACACAGTA CGATACCGTT TGTATTACA TGTCGTAGGT TATTTTAGC ACCTTAGTAT
F2
K T F S N G C D Y V S N K G M D T V S V G N T L Y Y V N K Q E G K S
1501 AAGACATTTT CTAACGGGTG CGATTATGTA TCAATAAAG GATGGACAC TGTTCTGTGA GGAACACAT TATATTATGT AATAAGCAA GAAGGTAAAA
TTCTGTAAAA GATTCGCCAC GCTATACAT AGTTTATTTT CCTACCTGTG ACACAGACAT CCATTTGTGA ATATAATACA TTTATTCGTT GTTCCATTTT
F2
L V V K G E P I I N F Y D P L V F P S D E F D A S I S Q V N E K I
1601 GTCTCTATGT AAAAGGTGAA CCAATAATAA ATTTCTATGA CCCATTAGTA TTCCCTCTG ATGAATTTGA TGCATCAATA TCPCAGTCA ACGAGAAGAT
CAGAGATACA TTTTCCACTT GGTATTATT TAAAGATACT GGGTAATCAT AAGGGGAGAC TACTTAAACT ACGTAGTTAT AGAGTTCACT TGCTTCTTA
F2
TM Domain
N Q S L A F I R K S D E L L H N V N A G K S T T N I M I T T I I I
1701 TAACCGAGGC CTAGCATTTA TTCGTAATC CGATGAATTA TTACATAATG TAAATGCTGG TAAATCCACC ACAAAATCA TGATACTAC TATAATTTA
ATTTGCTCTCG GATCGTAAAT AAGCATTTAG GCTACTTAAT AATGTATTAC ATTTACGACC ATTTAGGTGG TGTATTATGT ACTATTGATG ATATAATAT
TM Domain
Cytoplasmic Tail
V I I V I L L S L I A V G L L L Y C K A R S T P V T L S K D Q L S G
1801 GTGATTATAG TAATATTGTT ATCAATTAAT GCTGTTGGAC TGCTCTTATA CTGTAAAGCC AGAAGCACAC CAGTCACACT AAGCAAGAT CAACGTAGTG
CACTAATATC ATTATAACAA TAGTAATTA CGACAACCTG ACGAGAATAT GACATTCGGG TCTTCGTGTG GTCAGTGTGA TTCGTTTCTA GTTGACTCAC
FMDV 2A
membrane domain of WNVE (split)
Cytoplasmic Tail pre E/NS1 signal
I K N I A F S N N F D L L K L A G D V E S N P G P A R D R S I A L
1901 GTATAATAA TATTCGATTT AGTAACAAT TTGATCTGCT CAACTTGCA GCGATGTAG AATCAATCC TGGACCCGCC CGGGACAGGT CCATAGCTCT
CATATTTAT ATAACGTAAA TCATTTTAA AACTAGACGA GTTGAACG CCGCTACATC TTACTTTAGG ACCTGGGCGG GCCTGTCCA GGTATCGAGA
Transmembrane domain of WNVE (split)
NS1

```

· T F L A V G G V L L F L S V N V H A D T G C A I D I S R Q E L R
 2001 CACGTTTCTC GCAGTTGGAG GAGTTCTGCT CTCCTCTCC GTGAACGTGC ACECTGACAC TGGTGTGCC ATAGACATCA GCCGGCAAGA GCTGAGA
 GTGCAAGAG CGTCAACCTC CTCAGACGA GAAGGAGAGG CACTTGCACG TGCAGCTGTG ACCCACACGG TATCTGTAGT CGGCCGTTCT CGACTCT

PIV-WNV(ΔC)/RSV-F

1 GATCCTAATA CGACTCACTA TAGAGTAGTT CGCCTGTGTG AGCTGACAAA CTTAGTAGTG TTTGTGAGGA TTAACAACAA TTAACACAGT GCGAGCTGTT
 CTAGGATTAT GCTGAGTGTAT ATCTCATCAA GCGGACACAC TCGACTGTTT GAATCATCAC AAACACTCCT AATTGTTGTT AATTGTGTCA CGCTCGACAA
 N-terminus of C

 M S K K P G G P G K S R A V N M L K R G M P R V L S L
 101 TCTTAGCACG AAGATCTCGA TGTCTAAGAA ACCAGGAGGG CCGGCAAGA GCCGGGCTGT CAATATGCTA AAACGGCGAA TGCCCGCGT GTGTCTCTG
 AGAATCGTGC TTCTAGAGCT ACAGATTCTT TGGTCTCTCC GGGCCGTTCT CGGCCCGACA GTTATACGAT TTTGCCGCTT ACGGGCCGCA CAACAGGAAC
 N-terminus of C F signal

 NS3 cleavage F1

 I G L K Q K K R G G E L L I L K A N A I T T I L T A V T F C F A S G
 201 ATTGGACTTA AGCAAAAGAA GCGAGGGGGC GAGTTGCTAA TCCTCAAAGC AAATGCAATT ACCACAATCC TCACTGCAGT CACATTTTGT TTTGCTTCTG
 TAACTGAAAT TCGTTTTCTT CGCTCCCCCG CTCACAGATT AGGAGTTTCG TTTACGTTAA TGGTGTAGG AGTGACGTCA GTGTAAACA AAACAGGAC
 F1

 Q N I T E E F Y Q S T C S A V S K G Y L S A L R T G W Y T S V I T
 301 GTCAAAACAT CACTGAAGAA TTTTATCAAT CAACATGCAG TGCAGTTAGC AAAGGCTATC TTAGTGTCTT GAGAAGTGGT TGGTATACCA GTGTTATRAC
 CAGTTTTGTA GTGACTTCTT AAAATAGTTA GTGTACGTC ACGTCAATCG TTTCCGATAG AATCACGAGA CTCTTGACCA ACCATATGTT CACAATATTG
 F1

 I E L S N I K E N K C N G T D A K V K L I K Q E L D K Y K N A V T
 401 TATAGATTA AGTAATATCA AGGAAATTA GTGTATGGA ACAGATGCTA AGGTAAATTT GATAAACA GAATTAGATA AATAAATAA TGCTGTAACA
 ATAICTAAT TCATTATPAG TCCITTTTATT CACATTACCT TGCTACGAT TCCATTTTAA CPATTTTGTT CTTAATCTAT TTTATTTTTT ACGACATTGT
 F1

 E L Q L L M Q S T P P T N N R A R R E L P R F M N Y T L N N A K K T
 501 GAATTGCAGT TCCTCAAGCA AAGCACACCA CCAACAACA ATCGAGCCAG AAGAGAACCTA CCAAGTTTA TCAATTATAC ACTCAACAT GCCAAAAAAA
 CTTAACGTC ACGAGTACGT TTCGTGTGGT GTTGTTTGT TAGCTCGCTC TCTCTTGAT GGTCCAAAT ACTTAATATG TGAGTTGTTA CGGTTTTTTT
 F2

 F1

 N V T L S K K R K R R F L G F L L G V G S A I A S G V A V S K V L
 601 CCAATGTAAC ATTAACCAAG AAAAGGAAAA GAAGATTCTT TGGTTTTTTC TTAGGTGTTG GATCTGCAAT CGCCAGTGGC GTTGTGTAT CTAAGGTCTT
 GGTATACATG TAATTCTGTC TTTTCTTTTT CTCTAAAGA ACCAAAAAAC AATCCACAAC CTAGACGTTA GCGGTCAACC CAACGACATA GATTCCAGGA
 F2

 H L E G E V N K I K S A L L S T N K A V V S L S N G V S V L T S K
 701 GCACCTAGAA GGGGAAGTGA ACAAGATCAA AAGTCTCTA CTATCCACAA ACAAGGCTGT AGTCAGCTTA TCAATGGAG TTAGTGTCTT AACCCAGAAA
 CGTGGATCTT CCCCTTCACT TGTTCTAGTT TFCAGGAGAT GATAGGTTGT TGTCCGACA TCAGTCAAT AGTTTACCTC AATCACAGAA TTGCTCTTT
 F2

 V L D L K N Y I D K Q L L P I V N K Q S C S I S N I E T V I E F Q Q

801 GTGTTGAGCC TCAAAAAC TAATAGATAAA CAATTTGTTAC CTATTTGTGAA CAAGCAAAGC TGCAGCATAT CAAATATAGA AACTGTGATA GAGTCCCAAC
 CACAATCTGG AGTTTTTGGT ATATCTATTT GTTAACAATG GATAACACTT GTTCGTTTTG ACGTCTGATA GTTTATATCT TTGACACTAT CTCAAGGTTG
 F2

· K N N R L L E I T R E F S V N A G V T T P V S T Y M L T N S E L L ·
 901 AAAAGAACAA CAGACTACTA GAGATTACCA GGAATPTAG TGTTAATGCA GGTGTAACTA CACCTGTAAG CACTTACATG TTAACATAA GTCAATTATT
 TTTCTTCTTT GTCGTATGAT CTCTAATGGT CCCCTAAATC ACAATTACGT CCACATTGAT GTGGACATTC GTGAATGATC AATTGATTAT CACTTAATAA
 F2

· S L I N D M P I T N D Q K K L M S N N V Q I V R Q Q S Y S I M S I ·
 1001 GTCATTAATC AATGATATGC CTATAACAAA TGATCAGAAA AAGTTAATGT CCAACAATGT TCAATATGTT AGACAGCAAA GTTACTCTAT CATGTCCATA
 CAGTAAATG TIACATATCG GATATTGTTT ACTAGTCTTT TTCAAITACA GGTGTGTACA AGTTTATCAA TCTGTCTGTT CAATGAGATA GTACAGGTAT
 F2

· I K E E V L A Y V V Q L P L Y G V I D T P C W K L H T S P L C T T N ·
 1101 ATAAAAGAGG AAGTCTTAGC ATATGTAGTA CAATTACCAC TATATGGTGT TATAGATACA CCTGTGGA AACTACACAC ATCCCTCTA TGTACAACCA
 TATTTCTCC TCAGAAATCG TATACATCAT GTTAATGGTG ATATACCACA ATATCTATGT GGGACAACCT TTGATGTGTG TAGGGGAGAT ACATGTGTGT
 F2

· T K E G S N I C L T R T D R G W Y C D N A G S V S F F P Q A E T C ·
 1201 ACACAAAAGA AGGGTCCAAC ATCTGTITAA CAAGAAGTGA CAGAGAGTGG TACTGTGACA ATGCAGGATC AGTATCTTTC TTCCCAAGG CTGAAACATG
 TGTGTTTTCT TCCAGGTTG TAGACAATTT GTTCTTGACT GTCTCCTACC ATGACACTGT TACGCTCTAG TCATAGAAAG AAGGGTGTTC GACTTTGTAC
 F2

· K V Q S N R V F C D T M N S L T L P S E I N L C N V D I F N P K Y ·
 1301 TAAAGTTCAA TCAAATCGAG TATTTGTGA CACAATGAAC AGTTTACAT TACCAAGTGA AATAAATCTC TGCATGTG ACATATTCAA CCCCAATAT
 ATTTCAAGTT AGTTTAGCTC ATAAAACACT GTGTACTTGT TCAAATGTGA ATGGTTCACT TTATTTAGAG ACGTTACAAC TGATAAGTT GGGGTTTATA
 F2

· D C K I M T S K T D V S S S V I T S L G A I V S C Y G K T K C T A S ·
 1401 GATTGTAATA TTATGACTTC AAAACAGAT GTAAGCAGCT CCGTATACAC ATCTCTAGGA GCCATTGTGT CATGCTATGG CAAAACATAA TGTACAGCAT
 CTAACATTTT AATACTGAAG TTTTGTCTA CATTCTGCGA GGCAATAGTG TAGAGATCCP CGGTAACACA GTACGATACC GTTTGATTT ACAATGCGTA
 F2

· N K N R G I I K T F S N G C D Y V S N K G M D T V S V G N T L Y Y ·
 1501 CCAATAAAA TCGTGGAAATC ATAAAGACAT TTTCTAACGG GTGCGATTAT GTATCARATA AAGGGATGGA CACTGTGTCT GTAGTARCA CATTATATTA
 GGTATTTTT AGCACCTTAG TATTTCTGTA AAAGATTGCC CACGCTAATA CATAGTTTTT TTCCTACCT GTGACACAGA CATCCATTGT GTAATATAAT
 F2

· V N K Q E G K S L Y V K G E P I I N F Y D P L V F P S D E F D A S ·
 1601 TGAATATAG CAAGAAGGTA AAAGTCTCTA TGTAAGAGGT GAACCAATRA TAAATTTCTA TGACCCATAA GIATTCCTCT CTGATGATTT TGATGATCCA
 ACATTTATTC GTTCTCCAT TTTAGAGAT ACATTTTCCA CTCTGTTTAT ATTTAAGAT ACTGGTAAAT CATAGGGGA GACTACTTAA ACTACGTAGT
 F2

TM Domain
 ~
 1701 I S Q V N E K I N Q S L A F I R K S D E L L H N V N A G K S T T N I ·
 ATATCTCAAG TCAACGAGAA GATTAACCAAG AGCCTAGCAT TTATTCGTA ATCCGATGAA TTATTACATA ATGTAAATGC TGGTAAATCC ACCACAATA
 TATAGAGTTC ACTTGTCTTT CAAATGGTC TCGGATCGTA AATAAGCATT TAGGCTACTT AATAATGAT TACATTTAGC ACCATTAGG TGGTGTATT
 TM Domain

Cytoplasmic Tail
 ~
 · K I T T I I I V I I V I L L S L I A V G I L L Y C K A R S T P V T ·

1801 TCATGATAAC TACTATAATT ATAGTGATTA TAGTAATATT GTTATCATTG ATTGCTGTTG GACTGCTCCT ATACTGTAAAG GCCAGAAGCA CACCAGCTCAC
AGTACTATTG ATGATATATA TATCACTAAT ATCATTATAA CAATAGTATP TAACGACACAC CTGACGAGAA TATGACATTG CCGTCTTCGT GTGGTCAGTG
FMDV 2A

Cytoplasmic Tail

1901 · L S K D Q L S G T N N I A F S N N F D L L K L A G D V E S N P G P
ACTAAGCAA GATCAACTGA TGGTATAAA TAATATTGCA TTATAGTAACA ATTTTGATCT GCTCAAACCT GCAGGCGATG TAGAATCAAA TCCTGGACCC
TGATTCGTTT CTAGTTGACT CACCATATTT ATTATACGT AAATCATTTG TAAACTAGAA CGAGTTTGAA CGTCCGCTAC ATCTTAGTTT AGGACCTGGG
prM

C/prM signal

2001 G G K T G I A V M I G L I A C V G A V T L S N F Q G K V M M T V N A
GGAGAAAGA CCGGATATCC ACTCATGATT GGCCTGATCG CCTCCGTAGG AGCAGTTACC CTCTCTAACT TCCAAGGGAA GGTGATGATG ACGGTAAATG
CCTCCTTTCT GGCATAACG TCAGTACTAA CCGGACTAGC GGAGCGATCC TCGTCAATGG GAGAGATTGA AGGTTCCCTT CCACTACTAC TGCCATTAC
prM

2101 · T D V T D V I T I P T A A G K N L C I V R A M D V G Y M C D D T I
CTACTGACGT CACAGATGC ATCAGGATTC CAACAGCTGC TGGAAAGAAC CTAATGCATTTG TCAGAGCAAT GGATGTGGGA TACATGTGGG ATGATACTAT
GATGACTGCA GTGTCTACAG TAGTGTCAAG GTTGTGACAG ACCTTCTTTG GATACGTAAC AGTCTCGTTA CCTACACCCCT ATGTACACCC TACTATGATA
prM

2201 · T V E C P V L S A G N D P E D I D C W C T K S A V Y V R Y G R C T
CACTTAIGAA TGCCAGTGC TGTCCGCTGG TAATGATCCA GAAGACATCG ACTSTGGTG CACAAAGTCA GCAGTCTACG TCAGGTATGG AAGATGCACC
GTGAATACIT ACGGGTCACG ACGCCCGACC ATTACTAGGT CTTCGTAGC TGACAACCAC GTGTTTCAGT CGTCAGATGC AGTCCATACC TTCTACGTGG
prM

2301 K T R H S R R S R R S L T V Q T H G E S T L A N K K G A W M D S T K
AAGACACGCC ACTCAAGACG CAGTCGGAGG TCACTGACAG TCCAGACACA CCGAGAAAGC ACTCTAGCGA ACAAGAAGCG GGCTTGGATG GACAGCACCA
TTCTGTGGGG TGAATCTTGC GTACGCCCTCC AGTGACTGTC ACGTCTGIGT GCCTCTTTCC TGAGATCGGT TGTTCITCCC CCGAACCTAC CTGTCTGGT
prM

2401 · A T R Y L V K T E S W I L R N P G Y A L V A A V I G W M L G S N T
AGGCCACAC GTATTTGGTA AAAACAGAAT CATGGATCTT GAGGAACCCCT GGATATGCCC TGGTGGCAGC CGTCATTGCT TGGATGCTTG GGAGCAACAC
TCCGGTGTTC CATAAACCAT TTTTGTCTTA GTACCTAGAA CTCCCTGGGA CCTATACGGC ACCACCCCTG GCAGTAACCA ACCTACGAAC CCTCGTGTG
prM

2501 · M Q R V V F V V L L L L V A P A Y S F N C L G M S N R D F L E G V
CATGCAGAGA GTTGTGTTTG TCGTGTATT GCTTTTGGTG GCCCCAGCTT ACAGCTTTAA CTGCCCTGGA ATGAGCAACA GAGACTTCTT GGAAGGACTG
GTACGTCTCT CAACACAAAC AGCAGGATAA CGAAAACCAC CCGGGTCCGA TGTGAAATTT GACGGAACCT TACTGTTTGT CTCTGAAGAA CCTTCCTCAC
prM

E

2601 S G A T W V D L V L E G D S C V T I M S K D K P T I D V K M M N M E
TCTGGAGCAA CATCGGTGGA TTTGGTTCTC GAAGGGCACA GCTGCGTGAC TATCATGTCT AAGGACAAGC CTACCATCGA TGTGAGATG ATGATATGG
AGACCTCGIT GTACCCACCT AAACCAAGAG CTTCGGCTGT CGAGCGCATG APAGTACAGA TTCCTGTTCC GATGGTAGCT ACACCTCTAC TACTTATACC
E

2701 · A A N L A E V R S Y C Y L A T V S D L S T K A A C P A M G E A H N
AGGGCCCAA CTTGGCAGAG GTCCGCAGTT ATGCTATTTT GCCTACCCTC AGCGATCTCT CCACCAAAGC TGCGTGCCCG GCCATGGGAG AAGTCCAAA
TCCGCCGTTT GGACCTCTC CAGCCCTCAA TAACGATAAA CCGAGTGGAG TCGTAGAGA GGTGTTTTCG ACGCACGGGC CGTACCCTC TTCAGTGT
E

```

.....
2801 · D K R A D P A F V C R Q G V V D R G W G N G C G L F G K G S I D T
TGACAAAGCT GCTGACCCAG CTTTTCGTG CAGACAAGGA GTGGTGGACA GGGGCTGGGG CAACGGCTGC GGACTATTTG CCAAAGGAAG CATTGACACA
ACTGTTTGCA CGACTGGGTC GAAAACACAC GTCTGTTCTT CACCACCTGT CCCCAGACCC GTTGGCGAGC CCGTATAAAC CGTTTCTTC GTAACCTGTG
E
.....
2901 C A K F A C S T K A I G R T I L K E N I K Y E V A I F V H G P T T V ·
TGGCCAAAT TTGCTGCTC TACCAAGGCA ATAGGAAGAA CCATTTTGAA AGAGAAATATC AAGTACGAAG TGGCCATTTT TGTCATGGA CCAACTACTG
ACGCGGTTTA AACGGACGAG ATGGTTCCTT TATCCTTCTT GGTAAAACCT TCTCTTATAG TTCATGCTTC ACCGGTAAA ACAGGTACCT GGTGATGAC
E
.....
3001 · E S H G N Y S T Q V G A T Q A G R F S I T P A A P S Y T L K L G E ·
TGGAGTCGCA CGGAAACTAC TCCACACAGG TTGGAGCCAC TCAGGCAGGG AGATTACGCA TCACTCCTGC GCGCCCTCA TACACACTAA AGCTTGGAGA
ACCTCAGCGT GCCTTTGATG AGTGTGTCC AACCTCGGTG AGTCCGTCC TCTAAGTCTG ASTGAGGAGC CCGCGGAAGT ATGTGTGATT TCGAACCTCT
E
.....
3101 · Y G E V T V D C E P R S G I D T N A Y Y V M T V G T K T F L V H R ·
ATATGGAGG GTGACAGTGG ACTGTGAACC ACGGTGAGGG ATTGACACCA ATGCATACTA CGTGATGACT GTTGGAAACA AGACTTCTT GTCCATCTCT
TATACCTCTC CACTGTACCC TGACACTTGG TCCAGTCC TAACCTGTGT TACGTATGAT GCACACTGTA CAACCTTCTT TCTGCAAGAA CCAGTAGCA
E
.....
3201 E W F M D L N L P W S S A G S T V W R N R E T L M E F E E P H A T K ·
GAGTGTTC TGGACCTCAA CCTCCCTGG AGCAGTGTCT GAATACTCT GTGGAGGAC AGAGAGACCT TAATGGAGT TGAGGAACA CAGGCCAGCA
CTCACCAAGT ACCTGGAGT GAGGGAAAC TCGTCAGAC CTTCATGACA CACCTCCCTG TCTCTCTGCA ATTACCTCAA ACTCCTTGGT GTGGGTGCT
E
.....
3301 · Q S V I A L G S Q E G A L H Q A L A G A I P V E F S S N T V K L T ·
AGCAGTCTGT GATAGCATTG GGCTCACAAG AGGGACTCT GCATCAAGCT TTGGCTGGAG CCATTCCTGT GGAATTTCA AGCAACACTG TCAAGTTGAC
TGTTCAGACA CTATCTAAC CCGAGTGTTC TCCCTCGAGA CGTAGTTCGA AACCGACCTC GGTAAAGACA CCTTAAAAGT TCGTTGTGAC AGTTCAACTG
E
.....
3401 · S G H L K C R V K M E K L Q L K G T T Y G V C S K A F K F L G T P ·
CTCCGGTCAT TTGAAGTGA GAGTGAAGAT GGAATAATTG CAGTTGAAGG GAACAACCTA TGGCGTCTGT TCAAAGGCTT TCAAAGTTCT TGGGACTCCC
CAGCCAGTA AACTTCATAT CCACTTCA CTTTPTAAC GTCAACTTCC CTGTTGGAT ACCGCAGACA AGTTTCCGAA AGTTCAAAGA ACCCTGAGGG
E
.....
3501 A D T G H G T V V L E L O Y T G T D G P C K V P I S S V A S L N D L ·
GCAGACACAG GTCACGGCAC TGTGGTGTG GAATTGCAGT ACCTGGCAC GGATGGACCT TGCAAGTTC CTATCTGTC AGTGGCTTCA TTGAACGACC
CCTCTGTGTC CAGTCCCGT ACACCACAAC CTTAACGTCA TGTGACCGT CCTACCTGSA ACGTTTCAAG GATAGACGAG TCACCCGAGT AACTGTCTGG
E
.....
3601 · T P V G R L V T V N P F V S V A T A N A K V L I E L E P P F G D S ·
TAACGCCAGT GGGCAGATT GTCACGTCA ACCCTTTGT TTCAGTGGCC ACGGCCAAG CTAAGGTCTT GATTGAATTG TAACCACTT TGGAGACTC
ATTGCGGTCA CCCCTTAAC CAGTGACAGT TGGGAAAACA AAGTCACCGG TGCCGGTTC GATTCAGGA CTAACCTTAC CTTGGTGGGA AACCTCTGAG
E
.....
3701 · Y I V V G R G E Q Q I N H H W H K S G S S I G K A F T T T L K G A ·
ATACATAGT GTGGGACAG GAGAACACA GATCAATCAC CACTGGACA AGTCTGGAG CAGCATGGC AAAGCCTTA CAACCACTT CAAGGAGCG
TATGTATCAC CACCCGTCT CTTTGTGT CTAGTTAGT GTGACCGT CTAGACCTTC GTCCTAACG TTTGGAAAT GTTGGTGGGA GTTCTCTCG
E
.....
3801 Q R L A R L G D T A W D F G S V G G V F T S V G K A V H Q V F G G A ·
CAGAGACTAG CCCTCTAGG AGACACAGT TGGGACTTTG GATCAGTTG AGGGGTGTC ACCTCAGTTG GGAAGGCTT CCATCAAGT TCCGAGGAG
GTCTCTGATC GGGGAGATC TCTGTGCA ACCCTGAAC CTAGTCAACC TCCCAACAAG TGGAGTCAAC CCTTCCGACA GGTAGTTAC AAGCGTCTC

```

E

· F R S L F G G M S W I T Q G L L G A L L L W M G I N A R D R S I A ·
3901 CATTCCGCTC ACTGTCGGA GGCATGTCTT GGATAACGCA AGGATTGCTG GGGGCTCTCC TTTTGTGGAT GGGCATCAAT GCTCGTGACA GGTCATAGC
GTAGGCGGAG TGACAAGCCT CCGTACAGGA CCTATTGGGT TCCTAACGAC CCGCGAGAGG ACARACACCTA CCCGTAGTFA CGAGCACTGT CCAGGTATCG
NS1

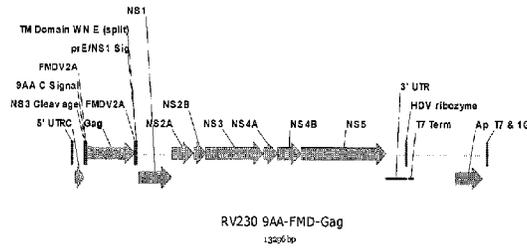
E

· L T F L A V G G V L L F L S V N V H A D T G C A I D I S R Q E L R
4001 TCTCACSTTT CTGCGAGTTG GAGGAGTICT GCTCTTCCTC TCCGCGAACG TGCACGCTGA CACTGGGTET GCCATAGACA TCAGCCGGCA AGAGCTGAGA
AGAGTGCAAA GAGCGTCAAC CTCTCAAGA CGAGAAGGAG AGGCACCTGC ACGTGCGACT GTGACCCACA CGGTATCTGT AGTCGGCCGT TCICGACTCT

Sequence Appendix 6

Construct 1

1. PIV-WN (Δ CprME)-SIV 9AA FMD Gag



2. Sequence of PIV-WN (Δ CprME)-SIV 9AA FMD Gag (partial).

C
~~~~~

5' UTR

-----

1 AGTAGTTCGC CTGTGTGAGC TGACAACTT AGTAGTGTG GTGAGGATTA ACAACAATTA ACACAGTGGC AGCTGTTTCT TAGCACGAAG ATCTCGAIGT M S  
TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCCTAAT TGTGTTAAT TGTGTACCGC TCGACAAAGA ATCGTGCTTC TAGAGTACA  
C

-----

101 K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K R A M L  
CTAAGAAACC AGGAGGGCCC GGCAGAGGCC GGGCTGCTA TTTGCTAATA CGCGGAATGC CCGCGCTGTT GTCCTTGATT GGACTTAAGA GGGCTATGTT  
GATTCCTTGG TCCTCCCGGG CCGTTCCTGG CCGCACAGAT AAACGATTTT GCGCCTTACG GGGCGCACAA CAGGAACATA CCGAATTCT CCGCATACAA  
C

-----

201 S L I D G K G P I R F V L A L L A F F R F T A I A P T R A V L D R  
GAGCCTGATC GACGSCAAGG GGCCAATACG ATTTGTGTTG GCTCTCTGG CGTTCTTCAG GTTCACAGCA ATTGCTCCGA CCCGAGCAGT GCTGGATCGA  
CTCGGACTAG CTGCCCTTCC CCGGTTATGC TAAACACAC ACAGAGAAACC GCAAGAAGTC CRAGTGTCGT TAACGAGGCT GGGCTCGTCA CGACCTAGCT  
C

-----

NS3 Cleavage  
~~~~~

301 W R G V N K Q T A M K H L L S F K K E L G T L T S A I N R R S S K Q
TGGAGAGGTG TGAACAACA AACAGCGATG AACACCTTC TGRGTTTCAA GAAGGAACATA GGGACCTTGA CCAGTGCTAT CAATCGCGGG AGCTCAAGC
ACCTCTCCAG ACTTGTGTTGT TTGTCGCTAC TTTGTGGAAG ACTCAAAGTT CTTCCTTGAT CCGTGGAACT GGTACAGATA GTTAGCCGCC TCGAGTTTTC
FMDV2A

9AA C Signal

```

-----
NS3 Cleavage
-----
          Gag
-----
401  · K K R G G K T G I A V I N F D L L K L A G D V E S N P G P M G A R ·
AGAAAAAGCG GGGCGGAAG ACAGGTATTG CTGTGATCAA TTTGACCTG TAAAACTGG CCGGGGACGT CGAAAGCAAC CCCGGTCCGA TGGCGCTAG
TCTTTTCGC CCCGCCTTC TGTCCATAAC GACACTAGTT AAAACTGGAC AATTTTGACC GGCCCTGCA GCTTTCGTTG GGGCCAGGCT ACCCGCGATC
          Gag
-----
501  · N S V L S G K K A D E L E K I R L R P G G K K K Y M L K H V V W A ·
GAATAGCGTG CTFAGTGGCA AAAAGGCIGA TGAAC TTGAG AAGATCGGEC TCGTCCGGG CGGGAAGAAG AAGTATATGT TGAACATGT CGTGTGGGCC
CTTATCGCAC GAATCACCGT TTTTCCGACT ACTTGAAC TCCTAGGCGG AGGCAGGCC GGCCTTCTTC TTCATATACA ACTTTGTACA GCACACCGGG
          Gag
-----
601  · A N E L D R F G L A E S L L E N K E G C Q K I L S V L A P L V P T G ·
GCCAARCAGT TAGATAGGTT TGGGCTAGCA GAGTCATTGC TCGAAAACAA GSAAGGATGT CAGAAGATAC TAAGTGTCTT GGCACCTTTG GTACCCACGG
CGGTTGCTCA ATCATATCAA ACCCGATCGT CTCAGTAAGC AGCTTTGTTT CTTTCTTACA GTCTTCTATG ATTCACAGGA CCGTGGAAAC CATGGGTGCC
          Gag
-----
701  · S E N L K S L Y N T V C V I W C I H A E E K V K H T E E A K Q I V ·
GGTCTGAGAA CTTAAAGAGT CTGTATAACA CTGTGTGGCT GATCTGGTGC ATTCACCGCG AAGAGAAAGT GAAGCACACC GAAGAAGCTA AGCAAATAGT
CCAGACTCTT GAATTTTCA GACATATTGT GACACACGCA CTAGACCACG TAAGTGCGGC TTCTCTTCA CTTGTGTGG CTTCCTCAT TCGTTTATCA
          Gag
-----
801  · Q R H L V V E T G T A E T M P K T S R P T A P S S G R G G N Y P V ·
GCAGAGACAT TTGGTCTGG AAACCGGGAC CGCCGAGACT ATGCCAAA AATCCCGTCC AACCGTCCA AGTAGTGGAA GAGGAGGTA CTACCCCGTT
CGTCTCTGTA AACACGACC ITTGGCCCTG CGGGCTCTGA TAGGGTTTT GTAGGGCAGG TTGGCGAGT TCATCACCTT CTCCTCCATT GATGGGGCAA
          Gag
-----
901  · Q Q I G C N Y V H L P L S P R T L N A W V K L I E E K K F G A E V V ·
CAGCAATCG GGGGAATTA CTGTCACTC CTTTGTGAC CAAGACCCT CABTGCATGG GTCAACTCA TCGAGGAAA GAAGTTCGGA GCGGAAGTGG
GTCGTTTAGC CCCCCTTAA GCACGTAGAG GGAACAGTG GTTCTGGGA GTTACGTACC CAGTTTGTAGT AGCTCCTTTT CTTCAAGCCT CGCCTTACC
          Gag
-----
1001 · P G F Q A L S E G C T P Y D I N Q M L N C V G D H Q A A M Q I I R ·
TCCCAGGGTT CCAGGCACTG AGTGAAGGTT GCATCCCTTA TGACATCAAC CAGATGCTTA ACTCGGTGCG CGACCATCAG GCCCGCATGC AGATTATTCG
AGGGTCCCAA GGTCCGTGC TCACTTCCCA CGTGAGGGAT ACTGTAGTTG GTCTACCAAT TGACGCAGCC GCTGGTAGTC CCGCGCTAGC TCTAATAAGC
          Gag
-----
1101 · D I I N E E A A D W D L Q H P Q P A P Q Q G Q L R E P S G S D I A ·
GGACATAATC AACGAGGAGG CTGCAGACTG GGATTTGCAG CACCCCCAAC CCGCCCTCA GCAAGGGCAG CTAAGGGAGC CTTCCGGCAG CGACATAGCT
CCTGTATTAG TTGCTCCTCC GACGTCGAC CCTAACCTG GTGGGGTTG GCGGGGAGT CSTTCCCTGC GATTCCTCG GAAGGCCCTG GCTGTATCGA
          Gag
-----
1201 · G T T S S V D E Q I Q W M Y R Q Q N P I P V G N I Y R R W I Q L G L ·
GGGACTACTA GCCTCGTGA TGACAGATT CABTGGATGT ACAGACAGCA GAATCCGATC CCCGTTGGCA ACATCTACCG GCGCTGGATT CAACTCGGAC
CCCTGATGAT CGAGGCCACT ACTTGTCTAA GTTACCTACA TGTCTGCTGT CTTAGGCTAG GGGCAACCGT TGTAGATGG GCGACCTAA GTTGGAGCTG
          Gag
-----
1301 · Q K C V R M Y N P T N I L D V K Q G P K E P F Q S Y V D R F Y K S ·
TTCAGAAGTG CGTCAGATG TACAACCCCA CCAATATTCT GGATGTGAAA CAGGGGCGA AAGAGCCCTT TCAATCCTAC GTCGACCGTT TCTACAAAAG
AAGTCTTAC GCAGTCTTAC ATGTTGGGGT GGTATAAGA CCTACACTTT GTCCCGGCT TTCTCGGAA AGTTAGGATG CAGCTGGCAA AGATGTTTT

```

Gag

1401 · L R A E Q T D A A V K N W M T Q T L L I Q N A N P D C K L V L K G
TCTACGGCC GAGCAGACCG ATGCCGCGT GAAGAAGTGG ATGACACAGA CGCTCTGAT ACAGAATCCT AACCCGTATT GTAACCTCGT GCTGAAGGGC
AGATGCCCGG CTCGTCTGGC TACGGCGTCA CTTCTTGACC TACTGTGTCT GCGAGGACTA TGTCTTACGA TTGGGACTAA CATTGAGCA CGACTTCCCG

Gag

1501 L G V N P T L E E M L T A C Q G V G G P G Q K A R L M A E A L K E A
TTAGGGGTAA ACCCAACGCT GGAAGAAATG TTAACCGCCT CCCAGGGAGT TGGTGGACCC GGACAGAAGG CCCGGCTAAT GGCCGAGGGC CTGAAAGAAG
AATCCCCATT TGGGTTGCGA CCTTCTTTAC AATTGGCGGA CGGTCCCTCA ACCACCTGGG CCTGTCTTCC GGGCCGATTA CCGGCTCCGC GACTTCTTC

Gag

1601 · L A P V P I P F A A A Q Q R G P R K P I K C W N C G K E G H S A K
CATTTGGTCC AGTACCATT CTTTGTGCT CCGCACACA GAGAGGTCCC CGTAAACCGA TCAATGCTG GACTGTGGG AAGGAGGGGC ACTCCGGTAA
GTACCCGAGG TCATGGGTA GAAACACGAC GGCCTGTTT CTCTCCAGG SCATTTGGCT AGTTTACGAC CTTGACACCC TTCCTCCCGG TGAGCGGATT

Gag

1701 · Q C R A P R R Q G C W K C G K M D H V M A K C P D R Q A G F L G L
ACAATGTGGA GCGCTAGAC GTCAGGGGTG TTGAAGTGT GGTAAATGG ACCAGTTAT GSCCAATGC CCCGACAGAC AAGCCGGGT CCTCCGGTAA
TGTACAGCT CCGGATCTG CAGTCCCGAC AACCTTCACA CCATTTTACC TGTGCAATA CCGTTCAGG GGGCTGTCTG TTCGCCCAA GGAGCCCAAT

Gag

1801 G P W G K K P R N F P M A Q V H Q G L T P T A P P E D P A V D L L K
GGGCTTGGG GAAAGAACCC CAGAACTTC CCAATGGCGC AAGTACACCA GGGCTGACC CCGACCGCCO CCCAGAGGA CCACGCGGTA GACCTTTGA
CCGGAAACCC CTTTTTCGG GTCTTTGAAG GGTTACCGCG TTATGTGCT CCGGACTGG GGCTGGGGG GGGGTCTCT GGTGCGCAT CTGGAAACT

Gag

1901 · N Y M Q L G K Q Q R E S R E K P Y K E V T E D L L H L N S L F G G
AAAATATAT GCAGCTGGG AAGCAGCAGC GCGAGAGTAG AGAGAAGCCC TACAAGGAGG TTACGGAGA TCTGTACAC CTTAATTGCT TATTTGGTGG
TTTTGATATA CGTCGACCC TTCGTCGTC CGCTCTCATC TCTTTCGGG ATGTCTCTC AATGCCTTCT AGACATCTG GAATTAAGCA ATAAACCAAC
TM Domain WN E (split)

FMDV2A

Gag pre/NS1 Sig

2001 · D Q N F D L L K L A G D V E S N P G P A R D R S I A L T F L A V G
TGATCAGAA TTGACCTGC TTAACCTTC TGGCAGCTT GAGTCAAATC CCGGCCCTGC CCGGGACAGG TCCATAGCTC TCACGTTTCT CCGAGTTGGA
ACTAGCTTA AGCTGGAGC AATTGAAAC ACCGCTGCAA CTAGTTTAG GCCCGGAGC GGCCCTCTCC AGTATTCGAG AGTGAARA GGTCAACT
TM Domain WN E (split)

NS1

2101 G V L L F L S V N V H A D T G C A I D I S R Q E L R C G S G V F I H
GGAGTCTGC TCTTCTCTC CGTGAACGTG CACGCTGACA CTGGGTGTC CATAGACATC AGCCGCCAAG AGCTGAGATG TGAAGTGA GTGTTCATC
CCTCAAGAG AGAAGGAG GCACTTGCAC GTCCGACTCT GACCCACAGC GTATCTGTAG TCGGCCGTC TCAGCTCTAC ACCTTCACCT CACAAGTATG

NS1

2201 · N D V E A W M D R Y K Y Y P E T P Q G L A K I I Q K A H K E G V C
ACAATGATG GAGGGCTGG ATGGACGGT ACAAGTATTA CCTGAAACG CCACAAGGCC TAGCCAGAT CATTCAGAA GCTCATAAG AGGAGTGTG
TGTACTACA CCTCCGACC TACCTGGCCA TGTTCATAAT GGCATTTGC GGTGTCCGG ATCGGTCTTA GTAGTCTTT CGAGTATTC TTCCTCAC

NS1

2301 · G L R S V S R L E H Q M W E A V K D E L N T L L K
CGGTACGA TCAGTTCCA GACTGGAGCA TCAATGTGG GAAGCAGTGA AGGAGGACT GAACACTCTT TTGAAG


```

301  W R G V N K Q T A M K H L L S F K K E L G T L T S A I N R R S S K Q .
TGGAGAGGTG TGAACAAACA AACAGCGATG AACACCTTC TGAGTTTCAA GAAGGAACTA GGGACCTTGA CCAAGTGTAT CAATCGGCGG AGCTCAAAGC
ACCTCTCCAC ACTTGTGTTG TTGTCGCTAC TTTGTGGAAG ACTCAAAGTT CTTCCCTGAT CCTCGGAACT GGTACAGATA GTTAGCCGCC TCGAGTTTCG
EMDV2A
SAA C Signal
NS3 Cleavage
Gag
401  K K R G G K T G I A V I N F D L L K L A G D V E S N P G P M G A R .
AGAAAAGCG GGGCGGAAG ACAGGTATTG CTGTGATCAA TTTTGCCTG TAAAACCTGG CCGGGGACGT C3AAGGCAAC CCGGTGCCGA TGGGCGCTAG
TCITTTTCCG CCGGCCPTTC TGTCATAAC GACACTAGTT ANAAGTGGAC AATTTTGACC GCGCCCTGCA GCTTTGCTTG GGCACAGGCT ACCGCGCATC
Gag
501  N S V L S G K K A D E L E K I R L R P G G K K K Y M L K H V V W A .
GATAGCGTG CTTAGTGGCA AAAAGGCTGA TGAACCTGAG AAGATCCGGC TCCGTCCGGG CCGGAAGAAG AAGTATATGT TGAACATGT CGTGTGGGCC
CTTATCGCAC GAATCACCGT TTTCCGACT ACTTGAACCT TTCTAGCCGG AGCGAGGCC GCCCTTCTTC TTCATATACA ACTTTGTACA GCACACCCGG
Gag
601  A N E L D R F G L A E S L L E N K E G C O K I L S V L A P L V P T G .
GCCAACGAGT TAGATAGGTT TGGCTAGCA GAGTCATIGC TCGAAAACAA SGAAGGATGT CAGAAGATAC TAAGTGTCTT GGCACCTTTG GTACCCACGG
CGGTGTCTCA ATCPATCCAA ACCCGATCGT CTCAGTAACG AGCTTTGTTT CCTTCTACA GTCTTCTATG ATTCACAGGA CCGTGGAAAC CATGGGTGCC
Gag
701  S E N L K S L Y N T V C V I W C I H A E E K V K H T E E A K Q I V .
GGTCTGAGAA CTTAAGAGCT CTGTATAACA CTGTGTGGGT GATCTGGTGC ATTCAGCGCG AAGAGAAAGT GAAGCACACC GAAGAAGCTA AGCAAATAGT
CCAGACTCTT GAATTTCTCA GACATATTGT GACACACGCA CTAGACCAGC TAAGTGGCGG TTCTCTTCA CTTGCTGTGG CTCTTCTGAT TCGTTTATCA
Gag
801  Q R H L V V E T G T A E T M P K T S R P T A P S S G R G G N Y P V .
CCAGAGACAT TTGTCCTGG AACCGGGGAC CGCCGAGACT ATGCCCAAAA CATCCCGTCC AACCGCTCCA AGTACTGGAA GAGGAGGTAA CTACCCCGTT
CGTCTCTGTA AACCAACACC TTTGGCCCTG CCGGCTCTGA TACGGGTTTT GTAGGGCAGG TTGGCGAGGT TCATCACCTT CTCTCCATT GATGGGGCAA
Gag
901  Q Q I G G N Y V H L P L S P R T L N A W V K L I E E K K F G A E V V .
CAGCAAAICG GGGGGAATTA CGTGCATCTC CCTTTGTCAC CAAGGACCT CAATGCATGG GTCRAACTCA TCGAGGAAA GAAGTTGGGA GCGGAAGTGG
CTCGTTTACG CCCCTTAA CTACGCTAGG GGAACAGTG GTTCTGGGA GTTACGTACC CAGTTTGAFT AGCTCCCTTT CTTCACGCT CCGCTTACC
Gag
1001 P G F Q A L S E G C T P Y D I N Q M L N C V G D H Q A A M Q I I R .
TCCCAGGSTT CCAGGCAGT ASIGAAGGST GCACTCCCTA TGACATCAAC CAGATGCTTA ACTGCCTCGG CGACCATCAG CCGCGGATC AGATTATTCC
AGGGTCCCAA GTCCTCTGAC TCACCTCCCA CGTAGGGAT ACTGTAGTTG GTCTAGCAAT TGACGCAGCC GCTGTGATGC CCGGCTACG TCTAATAAGC
Gag
1101 D I I N E E A A D W D L Q H P Q P A P Q Q G Q L R E P S G S D I A .
GGACATATC AACGAGGAG CTGCAGACTG GGATTTGAG CACCCCAAC CCGCCCTCA GCAAGGGCAG CTAAGGGAGC CTTCGGGAG CGACATAGCT
CCTGTATTAG TTGCTCCTC GACGTCTGAC CCTAAACGTC GTGGGGTGG GCGGGGAGT CGTTCCTG GATTCCCTCG GAAGCCCTC GCTGTATCGA
Gag
1201 G T T S S V D E Q I Q W M Y R Q Q N P I P V G N I Y R R W I O L G L .
GGGACTACTA GCTCCGTGGA TGAACAGATT CAATGGATCT ACACAGAGCA GAATCCGATC CCGTGGCA ACATCTACCG GCGCTGATT CACTCCGAC
CCCTGATGAT CGAGCCACT ACTTGTCTAA GTTACCIACA TGCTGTCTGT CTTAGGCTAG GGGCAACCGT TGTAGATGCC CCGACCTAA GTTAGCCCTG

```

Gag

1301 · Q K C V R M Y N P T N I L D V K Q G P K E P F Q S Y V D R F Y K S ·
 TTCAGAAGTG CGTCAGAATG TACAACCCCA CCAATATTCT GGATGTGAAA CAGGGGCGGA AAGAGCCCTT TCAATCCTAC GTCGACCGTT TCTACAAAAG
 AAGTCTTAC GGAGTCTTAC ATCTTGGGGT GGTATAAGA CCTACACTTT GTCGCCGGCT TTCTCGGGAA AGTTAGGATG CAGCTGCCAA AGATGTTTTC
 Gag

1401 · L R A E Q T D A A V K N W M T Q T L L I Q N A N E D C K L V L K G ·
 TCTACCGCC GAGCAGACG ATGCCGAGT GAAGACTGG ATGACACAGA CGCTCTGAT ACAGANTGCT AACCTGATT GTAACCTGT GCTGAGGGC
 AGATGCGCG CTCCTTGGC TACCGCGTCA CTTCTTGACC TACTGTSTCT GCGAGGACTA TGTCTTACGA TTGGGACTAA CATTGAGCA CGACTTCCCG
 Gag

1501 · L G V N P T L E E M L T A C Q G V G G P G Q K A R L M A E A L K E A ·
 TTAGGGTAA ACCCAACGCT GGAAGAAATG ITAACCGCTT GCCAGGGAGT TGGTGGACCC GGACAGAAGG CCCGGCTAAT GGCCGAGGCG CTGAAAGAAG
 AATCCCAAT TGGGTGGCA CCTCTTTAC AATTGGCGGA CGGTCCCTCA ACCACCTGGG CTTGCTTCC GGGCCGATTA CCGGCTCCCG GACTTTCTTC
 Gag

1601 · L A P V P I P F A A A Q Q R G P R K P T K C W N C G K E G H S A K ·
 CATTGGCTCC AGTACCCATT CCTTTTGGCTG CCGCACAACA GAGAGGTCC CGTAAACCGA TCAATGCTG GAACGTGGG AAGGAGGGGC ACTCCGCTAA
 GTACCCGAGG TCAATGGTAA GGAACAACGAC GGCCTGTTGT CTTCCAGGG GCATTTGGCT AGTTTACGAC CTTGACACCC TTCTCCCCCG TGAGGCGATT
 Gag

1701 · Q C R A P R R Q G C W K C G K M D H V M A K C P D R Q A G F L G L ·
 ACAATGTCGA GCGCCTAGAC GTCAGGGGTG TTGGAAGTGT GGTAAATG ACCACCTTAT GGCCAAATGC CCCGACAGAC AAGCCGGGTT CCTCGGGTTA
 TGTACAGCT CCGGGATCTG CAGTCCCAC AACCTTACA CCATTTTACC TGGTGAATA CCGGTTTACG GGGCTGTCTG TTCGGCCCAA GGAGCCCAAT
 Gag

1801 · G P W G K K E R N F P M A Q V H Q G L T P T A P P E D P A V D L L K ·
 GGGCCTTGGG GAAAAAGCC CAGAACTTC CCAATGGCGC AAGTACACCA GGGCCTGACC CCGACCGCC CCCAGAGGA CCCAGCCGTA GACCTTGA
 CCGGAACCC CTTTTTCGG GTCTTGAAG GGTACCAGG TTAGTGGGT CCGGACTGG GCGTGGCGG GGGTCTCTT GGTGGCAT CTGAGAACT
 Gag

1901 · N Y M Q L G K Q R E S R E K P Y K E V T E D L L H L N S L F G G ·
 AAACTATAT GCAGCTGGG AAGCAGCAG CCGAGAGTAG AGAGAGCCC TACAAGGAGG TTACGGAAGA TCTGTACAC CTTAATTCGT TATTTGGTG
 TTTTATATA CGTCGACCC TTCGTCGTCG CGCTCTCATC TCTCTCGGG ATGTCTCC ACCGCTTCT AGACAATGT GAATTAAGCA ATAAACACC
 FMDV2A

Gag Pro

2001 · D Q N F D L L K L A G D V E S N P G P V L E L R Q R G P Q R Q A V ·
 TGATCAGAT TTCACCTGC TTAACCTGC TGGCGACGTT GAGTCAATC CCGGCCCTGT GCTGGAGTG AGACAGCGG GCGCCACGG CCAGGCTGT
 ACTAGTCTA AAGTGGAGC AATTTCAAC ACCGCTGCAA CTCAGTTTAC GCGCGGACA CGACCTCAAC TCTGTGCGG CCGGGTCCG CSTCCGACA
 Pro

2101 · Q S P S E T G L L E V W Q D G P R D G Q M P R Q T G G F F R P W S M ·
 CAGAGCCAT CAGACCGGG TCTACTTGAG GTGTGGCAGG ATGGCCCCG TGATGGACAG ATGCTCGCC AGACGGGAGG GTTCTCCGA CCTGGAGTA
 GTCTCGGTA GTCTGTGCC AGATCAACT CACACCGTCC TACCGGGGG ACTACCTGTC TACCGAGCGG TCTGCCCTCC CAAGAAGGCT GGGACCTCAT
 Pro

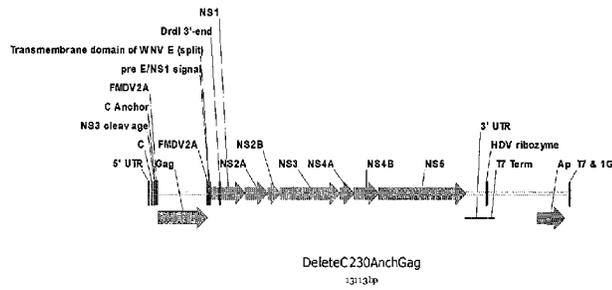
2201 · G K E A P Q F P H G S S A S G A D A N C S P R G P S C G S A K E L ·
 TGGSAAGGA GGCCCCGAG TTCCCTCATG GCTCTTCTGC CTCTGGCGG GATGCCAAT GTAGCCCGG AGGCCCTTCT TGCGGCTCAG CCAGGAGCT
 ACCCTTCTCT CCGGGCGTC AAGGAGTAC CGAGAAGAGG SAGACCCGG CTACGGTAA CATGGGGG TCCGGGAAG ACGCCGACT GTTCTCTCA
 Pro

```

.....
· H A V G Q A A E R K Q R E A L Q G G D R G F A A P Q F S L W R R P
2301 GCACGCAGTG GCCCAGGCGG CAGACGGCAA ACAGCGAGAA GCACGCGAGG CCGGTGACCC TGSTTTGCC CCCCACAAT TCAGTCTGTG GCGCCGACCT
CFTGCGTCAO CCGTCCGCTG GTCTCGCGTT TGTCGCTCTT CBTGACCTCC CGCCACTGGC ACCAAAACCG CCGGGTGTTA AGTCAGACAC GCGGGCTGGA
Pro
V V T A H I E G Q P V E V L L D T G A D D S I V T G I E L G F H Y T
2401 GTCCGTACTG CTCATATCGA GGGTCAGCCC GTGGAGGTTT TACTGGACAC TGGCGCAGAC GATTCTATTG TGACTGGCAT TGAAC TAGGC CCCCATTACA
CAGCACTGAC GAGTATAGCT CCCAGTGGG CACCTCCAAA ATGACCTGTG ACCCGCTCTG CTAAGATAAC ACTGACCGTA ACTTGATCCG GGGGTAATGT
Pro
· P K I V G G I G G F I N T K E Y K N V E I E V L G K R I K G T I M
2501 CTCCAAAAAT CGTAGGGGGG ATAGGAGGAT TTATCAACAC GAAGGAGTAT AAGAATGTGG AGATCCAGGT TCTCGGAAA CGCATTAAGG GAACGATTAT
GAGGTTTTTA GCATCCCCC TATCTCCTA AATAGTTGTG CTTCCTCATA TTCTTACACC TCTAGCTCCA AGAGCCTTTT GCGTAATTC CTTGCTAATA
FMDV2A
Pro
· T G D T P J N I F G R N L L T A L G M S L N L N F D L I K L A G D
2601 GACAGGCGAT ACACCCGATTA ACATCTTTGG ACGCAATCTA CTACGGCCC TCGGAATGAG CCTTAACCTC AACTCGACT TACTCAAGCT GCOCGGAGAC
CTGTCGCGTA TGTGGGTAT TGTAGAACC TGCGTTAGAT GAATCCCGGG AGCTPCTC GGAATGGAG TTGAAGCTGA ATGAGTTCGA GCGGCTCTCG
TM domain WN E (split)
pre/NS1 Sig
FMDV2A NS1
V E S N E G P A R D R S I A L T F L A V G G V L L F L S V N V H A D
2701 GTGGAGTCCA ATCCCGGCC AGCCCGGGAC ASGTCCATAG CTCTACGTT TCTCGCAGTT GGAGGAGTTC TGCTCTCCT CTCCGTGAC GTGCACGCTG
CACCTCAGST TAGGGCCGGG TCGGGCCCTG TCCAGGTATC GAGAGTGCAA AGAGCGTCAA CCTCCTCAG ACGAGAAGGA GAGGCATTG CACGTGCGAC
NS1
· T G C A I D I S R Q E L R C G S G V F I H N D V E A W M D R Y K Y
2801 ACACTGGGTG TGCCATAGAC ATCAGCCGGC AAGAGCTGAG ATGTGGAAGT GGAGTCTCA TACACAATGA TGTGGAGGCT TGGATGGACC GGTACAAGTA
TGTGACCCAC ACGGTATCTG TAGTCGGCCG TTCTCGACTC TACACCTTCA CCTCACAGT ATGTGTTACT ACACCTCCGA ACCTACCTGG CCAITGTCAT
NS1
· Y P E T P Q G L A K I I Q K A H K E G V C G L R S V S R L E H Q M
2901 TTACCTGAA ACGCCACAAG GCCTAGCCAA GATCATTGAG AAAGCTCATA AAGAGGAGT GTGCGGTCTA CGATCAGTTT CCAGACTCGA GCATCAATG
AATGGGACTT TGGGTGTTC CGGATCGGTT CTAGTAAGTC TTTCGAGTAT TCCTTCCCTA CACGCCAGAT GCTAGTCAA GGTCTGACCT CGTAGTTTAC
NS1
W E A V K D E L N F L L K
3001 TGGGAAGCAG TGAAGACGA GCTGAACACT CTTTTCAG
ACCCCTGCTC ACTTCTGCT CGACTTGGA GAAAACCTC

```

Construct 3
1. PIV-WN (Δ CprME)-SIV Anch Gag



2. Sequence of PIV-WN (Δ CprME)-SIV Anch Gag (partial).

```

                    C
                    -----
                    5' UTR
                    -----
1  AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTTC GTGAGGATTA ACAACAATTA ACACAGTGGC AGCTGTTTCT TAGCACGAAG ATCTCGATGT
   TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTGTGTAAT TGTGTCACGC TCGACAAAGA ATCGTGTCTC TAGAGCTACA
                    C
                    -----
                    NS3 cleavage
                    -----
101  K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R
     CTAAGAAACC AGGAGGCCCC GCAAGAGCC GGGCTGTCTA TTGCTAAAA CGCGGAATGC CCCGCGTGTG GTCCTTGATT GGACTTAAGC AAAAGAGCG
     GATTCTTTGG TCCTCCCGGG CCGTTCCTCG CCCGACAGAT AAACGATTTT GCGCCTTAGG GGGCGCACAA CAGGAACTAA CCTGAATTCG TTTCTTCGC
                    C Anchor
                    -----
NS3 cleavage
                    -----
201  G G K T G I A V M I G L I A S V G A N F D L L K L A G D V E S N P
     AGGCGGAAAG ACAGGTATTG CTGTGATGAT TGGCCTGATC GCCACCGTAG GAGCAAATTT TGACCTGTTA AAAGTGGCCG GGGACGTCGA AAGCAACCCC
     TCCGCTTTC TGTCCATAAC GACTACTA ACCGGACTAG CCGTCCGATC CTCGTTTAAA ACTGGACAAT TTGACCCGGC CCCTGCAGCT TTCGTTGGGG
     FMDV2A
                    -----
                    Gag
                    -----
301  G P M G A R N S V L S G K K A D E L E K I R L R P G G K K K Y M L K
     GTCGATGC GCGTAGGAA TAGCGTCTT AGTGGCAAAA AGCGTATGA ACTTGAGAAG ATCCGGCTCC GTCCGGCCGG GAAGAAGAAG TATATGTTGA
    
```

CCAGGCTACC CCGGATCCTT ATCCGACGAA TCACGCTTTT TCCGACTACT TGAACCTTTC TAGGCCGAGG CAGGCCGCGC CTTCCTTCTC ATATACAAC
 Gag

401 · H V V W A A N E L D R F G L A E S L L E N K E G C Q K I L S V L A
 AACATGTCGT CTGGGCGGCC AACGAGTTAG ATAGTITGG GCTAGCAGAG TCATTGCTCG AABACAAGGA AGGATGTCAG AAGATACTAA GTGCCTGGC
 TTGTACAGCA CACCCGGGGG TTGCTCAATC TATCCAAACC CGATCGTCTC AGTAACGAGC TTTTGTTCCT TCCTACAGTC TTCTATGATT CACAGGAGCC

Gag

501 · P L V P T G S E K L K S L Y N T V C V I W C I H A E E K V K H T E
 ACCTTTGGTA CCCACGGGGT CTGAGAACTT AAAGAGTCTG TATAACACTG TGTGGGTGAT CTGGTGCATT CAGCCCGAAG ACAAGTGA GACACCCGAA
 TGGAAACCAT GGGTGCCCA GACTCTTGAA TTTCTCAGAC ATATTGTGAC ACACGCACTA GACCACGTAA GTGGGGCTTC TCTTCACTT CGTGTGGCTT

Gag

601 · E A K Q I V Q R E L V V E T G T A E T M P K T S R P T A P S S G R G
 GAAGCTAAGC AAATAGTGCA GAGACATTG TGCCTGGAAA CCGGGACCGC CGAGACTATG CCCAAACAT CCCGTCCAAC CGCTCCAAGT AGTGAAGAG
 CTTCGATTGC TTTATCACGT CTCTGTAAC CAGCACCTTT GGCCCTGGCG GCTCTGATAC GGGTTTGTG GGGCAGGTTG GCGAGGTICA TCACCTTCTC

Gag

701 · G N Y P V Q Q I G G N Y V H I P L S P R T L N A W V K L I E E K K
 GAGGTAECTA CCCCGTTCAG CAAATCGGGG GGAATTACGT GCATCTCCCT TTGTCACCAA GAACCCCTCAA TGCATGGGTC AAATCATCG AGGAAAGAA
 CTCCATTGAT GGGCAAGTC GTTTAGCCCC CCTTAATGCA CGTAGAGGA AACAGTGGTT CTTGGGAGTT ACGTACCCAG TTTGAGTAC TCCTTTTCTT

Gag

801 · F G A E V V P G F Q A L S E G C T P Y D I N Q M L N C V G D H Q A
 GTTCGGAGCG GAAGTGSTCC CAGGGTTCCA GGCACCTGAT GAAGGGTGA CTCCTATGA CATCAACCAG ATGCTTAACT GCGTCGGCGA CCATCAGGCC
 CAAGCCTCGC CTTACCCAGG GTCCCAAGGT CCGTGACTCA CTTCCACGTT GAGGGATACT GTAGTTGGTC TACGAATTGA CCGAGCCGCT GGTACTCCGG

Gag

901 · A M Q I I R D I I N E E A A D W D L Q H P Q P A P Q Q G O L R E P S
 GCGATGAGA TTATTCGGGA CATAATCAAC GAGGAGGCTG CAGACTGGGA TTGTCAGCAC CCCCAACCG CCCCTCAGCA AGGGCAGCIA AGGGAGCCTT
 CGCTACGTCT AATAAGCCCT GTATTAGTTG CTCCCTCCGAC GTCTGACCT AAACGTCGTG GGGGTTGGG GGGGAGTCGT TCCGTCGAT TCCCTCGGAA

Gag

1001 · G S D I A G T T S S V D E Q I Q W M Y R Q Q N P I P V G N I Y R R
 CCGGACGGA CATAGCTGGG ACTACTAGCT CCGTGGATGA ACAGATCAA TGGATGTACA GACAGCAGAA TCCGATCCCC GTGGGCAACA TCTACCGGCG
 GCGCCCTCGT GTATCGACCC TGATGATCGA GGCACCTACT TGTCTAAGTT ACCTACATGT CTGTGCTCTT AGGCTAGGGG CAACCGTTG AGATGGCCCG

Gag

1101 · W I Q L G L Q K C V R M Y N P T N I L D V K Q G P K E P F Q S Y V
 CTGCATTCAA CTCGGACTTC AGAAGTGCCT CAGAATGATC AACCCACCA ATATTCTGGA TGTGAAACAG GGGCCGAAAG AGCCCTTICA ATCCTACGTC
 GACCTAAGTT GAGCCTGAAG TCTTACGCA GTCTTACATG TTGGGGTGGT TATAAGACCT ACACITGTC CCCGGCTTC TCGGAAAGT TAGGATGCA

Gag

1201 · D R F Y K S L R A E Q T D A A V K N W M T Q T L L I Q N A N P D C K
 GACCGTITCT ACAAAAGTCT ACGGCGCGAG CAGACCGATG CCGCAGTGA GAACTGGATG ACACAGACCC TCCTGATACA GAATGCTAAC CCTCATGTA
 CTGGCAARGA TCTTTTCAGA TGCCGGGCTC GTCGCTAC GCGCTACTT CTTGACCTAC TGTGTCTGCG AGGACIATGT CTTACGATTG GGACTAACAT

Gag

1301 · L V L K G L G V N P T L E E M L T A C Q G V G G P G Q K A R L M A
 AACTCGTGT GAAGGGCITA GGGGTAAACC CAACCGTGA AGAAATGTTA ACGGCTGCC AGGAGITGG TGGACCCGA CAGAAGGCC GGCTAATGG
 TTGAGCACGA CTCCCGAAT CCCGATTGG GTTCCGACCT TCTTTACAAT TGCCGGACCG TCCCTCAACC ACCTGGGCT GTCTCCGGG CCGATTACCG

Gag

```

1401 · E A L K E A L A P V P I P F A A A Q Q R G P R K P I K C W N C G K
CGAGCGCTG AAGAAGCAT TGCCTCCAGT ACCCATTCCT TTTGCTGCGG CACAACAGAG AGGTCCCGGT AAACCGATCA AATGCTGGAA CTGTGGGAAG
GTCGCCGAC TTTCTCGTA ACCGAGGTCA TGGTAAGGA AAACGACGGC CTGTTGCTC TCCAGGGCA TTTGGCTAGT TTACGACCTT GACACCTTC

-----
Gag
-----
1501 · E G H S A K Q C R A P R R Q G C W K C G K M D H V M A K C P D R Q A
GAGGGCACT CCGTAACA ATGTCGAGCG CTTAGACGTC AGGGGTGTTG GAAGTGTGGT AAAATGGACC ACGTTATGGC CAAATGCCCC GACAGACAAG
CTCCCCGTA GCGATTGT TACAGCTGCG GGATCTGCAG TCCCCACAC CTTACACCA TTTTACCTGG TCAATACCG GTTTACGGGG CTGTCTGTTC

-----
Gag
-----
1601 · G E L G L G P W S K K P R N F P M A Q V H Q G L T P T A P P E D P
CCGGTTCCT CCGTTAGGG CTTGGGAA AAAAGCCAG AAACCTCCCA ATGGCCGAAG TACACCGGG CCTGACCCCG ACCGCCCCCG CAGAGGACCC
GGCCCAAGGA GCCCAATCC GGAACCCCTT TTTTCGGGTC TTTGAAGGGT TACCGGCTTC ATGTGGTCCC GGACTGGGGG TGGCGGGGGG GTTCTCTGGG

-----
Gag
-----
1701 · A V D L L K N Y M Q L G K Q Q R E S R E K P Y K E V T E D L L H L
AGCCGTAGC CTCTTGA AAA ACTATATGCA GCTGGGGAAG CAGCAGCGCG AGAGTAGAGA GAAGCCCTAC AAGGAGGTA CCGAAGATCT GTTACACCTT
TCGGCATCTG GAGAACCTTT TGATATACGT CGACCCCTTC CTCGTCGCGC TCTCATCTCT CTTGGGGATG TTCCTCCAAT GCCTTCTAGA CAATGTGGAA

-----
FMDV2
pre E/NS1 signal
-----
Gag Transmembrane domain of WNV E
(split)
-----
1801 · N S L F G G D Q N F D L L K L A G D V E S N P G P A R D R S I A L T
AATTCGTAT TGGTGGTGA TCAGAATTC GACCTGCTTA AACTTGTCTG CGACGTGAG TCAATCCGG GCCCTGCCCG GGACAGTCC ATAGCTCTCA
TTAAGCAATA AACCCACT AGTCTTAAG CTGGACCAAT TTGAACGACC GCTGCAACT AGTTTAGGCC CGGACGGCG CCTGTCCAGG TATCGAGAGT

-----
Transmembrane domain of WNV E (split)
-----
NS1
-----
1901 · F L A V S G V L L F L S V N V H A D T G C A I D I S R O E L R C G
CGTTCTCGC APTTGAGGA GTTCTGCTCT TCTCTCCGT GAACSTGCAC GCTGACACTG GGTGTGCCAT AGACATCAGC CGGCAAGAGC TGAGATGTGG
GCAAGAGCG TCAACCTCT CAAGACGAGA AGGAGAGGCA CTTGCACGTG CGACTGTGAC CCACACGGTA TCTGTAGTGG GCCCTCTCG ACTCTACCC

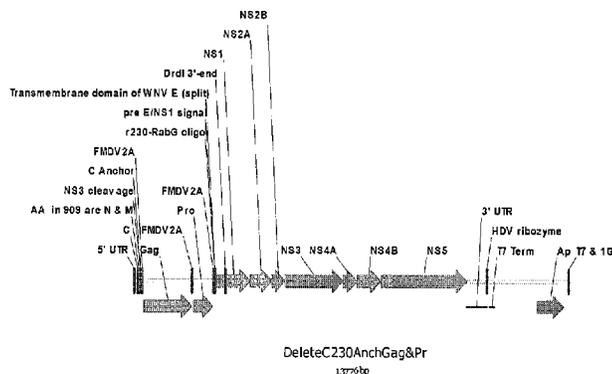
-----
NS1
-----
2301 · S G V F I H N D V E A W M D R Y K Y Y P E T P Q G L A K I I Q K A
AAGTGGAGT TGCATACACA ATGATGTGGA GGCCTGGATG CACCGGTACA AGTATTACCC TCAAAACGCCA CAAGGCCTAG CCAAGATCAT TCAGAAAGCT
TTCACCTCAC AAGTATGTET TACTACACCT CGAACCTAC CTCGCCATGT TCATATGGG ACTTTGGGT GTCCGGATC GGTCTAGTA AGTCTITCGA

-----
NS1
-----
2101 · H K E G V C G L R S V S R L E H Q M W E A V K D E L N T L L K
CATAGGAAG GAGTCTCGG TCTACGATCA GTTCCGAC TGGAGCATCA AATGTGGAA GCAGTGAAG ACGAGCTGAA CACTCTTTG AAG
GTATTCCTC CTCACACGCC AGATGCTAGT CAAAGGTCTG ACCTCSTAGT TTACACCTT GCCTACTTC TGCTCGACTT GTGAGAAAC TTC

```

Construct 4

1. PIV-WN (Δ CprME)-SIV Anch Gag & Pro



2. Sequence of PIV-WN (Δ CprME)-SIV Anch Gag & Pro (partial).

```

C
-----
5' UTR
-----
1 AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTTC GTGAGGATTA ACAACAATTA ACACAGTGC GCTGTTTCT TAGCACGAAG ATCTCGATGT
  TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTGTTTAA TGTGTCAAGC TCGACAAAGA ATCGTGCTC TAGAGCTACA
                                     NS3 cleavage
C
-----
  K K P G G P G K S R A V Y L L K R G M P R V I S L I G L K Q K K R
101 CTAGAAACC AGGAGGGCCC GGCAAGAGCC GGGCTGTCTA TTTGCTAAAA CGCGAATGC CCCCGCTGT GTCCCTGATT GGACTAAGC AAAAGAGCG
  GATTCTTTGG TCCTCCCGG CCSTTCTCCG CCCGACAGAT AATCGATTTT GGGCCTTACG GGGCGCACAA CAGGAACATA CGTGAATTCG TTTTCTTCG
                                     C Anchor
-----
NS3 cleavage
-----
  G G K T G I A V M I G L I A S V G A N F D L L K L A G D V E S N P
201 AGGCGGAAAG ACAGGTATTG CTGTGATGAT TGGCCTGATC GCCAGCGTAG GAGCAAATTT TGACCTGTTA AAACCTGGCG GGGACGTCGA AAGCAACCCC
  TCCGCTTTC TGTCCATAAC GACACTACTA ACCGGACTAG CGGTGCAATC CTCGTTTAAA ACTGGACAAT TTGACCGGC CCTGCAGCT TTCGTTGGG
FMDV2A
-----

```

```

-----
                                Gag
-----
301  G P M G A R N S V L S G K K A D E L E K I R L R P G G K K K Y M L K
GGTCCGATGG GCGCTAGGAA TAGCGTGCTT AGTGGCAAAA AGGCTGATGA ACTTGAGAAG ATCCGGCTCC GTCCGGGCGG GAAGAAGAAG TATATGTTGA
CCAGGCTACC GCGGATCCCT ATCGCACGAA TCACCGTTTT TCCGACTACT TGAAGCTTTC TAGGCCGAGG CAGGCCGCGC CTTCCTTCTC ATATACAACT
                                Gag
-----
401  H V V W A A N E L D R F G L A E S L L E N K E G C Q K I L S V L A
AACATGTGCT GTGGGCGCGC AACGAGTTAG ATAGGTTTGG GCTAGCAGAG TCATTGCTCG AAAACAAGGA AGGATGTCAG AAGATACTAA GTCTCCTGGC
TTGTACACCA CACCGGCGCG TTGCTCAATC TATCCAACC CGATCGCTCC AGTAACGAGC TTTTSTTCTC TCCTACAGTC TTCTATGATT CACAGGACCG
                                Gag
-----
501  P L V P T G S E N L K S L Y N T V C V E W C I H A E E K V K H T E
ACCTTTTGSTA CCCACGGGGT CTGAGAACTT AAAGAGTCTG TATAACACTG TGTGCGTGAT CTGGTGCATT CACGCCGAAG AGAAAGTGAA GCACACCGAA
TSGAAACCAT GGGTGCOCOA GACTCTTGAA TTTCTCACAC ATATTGTGAC ACACGCACTA GACCACGTAA GTGGGGGCTC TCTTTCACCT CGTGTGCTT
                                Gag
-----
601  E A K Q I V Q R H L V V E T G T A E T M P K T S R P T A P S S G R G
GAAGCTAAGC AAATAGTCCA GACACATTTG GTCGTGGAAA CCGGGACCGC CGAGACTATG CCCAAAACAT CCCGTCCAAC CGCTCCAAGT AGTGAAGAG
CTTCGATTCC TTTATCACCT CTCTGTAAC CAGCACCTTT GGCCCTGGCG GCTCTGATAC GGCTTTTGTA GGGCAGGTTG GCGAGATTCA TCACCTTCTC
                                Gag
-----
701  G N Y P V Q Q I G G N Y V H L P L S P R T L N A W V K L I E E K K
GAGGTAAC TA CCCGTTTCAG CAAATCGGGG GGAATTACGT GCATCTCCCT TTGTCAACCA GGACCCCTCAA TGCATGGGTC AAATCATCG AGGAAAAGAA
CTCCATTGAT GGGCAAGTC GTTTAGCCCC CCTTAATGCA CGTAGAGGGA AACAGTGGTT CCTGGGAGTT AGTACCAG TTTGAGTAGC TCCTTTTCTT
                                Gag
-----
801  F G A E V V P G F Q A L S E G C T E Y D I N Q M L N C V G D H Q A
GTTCCGAGCG GAAGTGGTCC CAGGGTTCCA GGCAGTGTG GAAGGTGCA CTCCCTATGA CATCAACCAG ATGCTTAACT CCGTCCGGCA CCATCAGGCC
CAGCCCTCC CTTCACCCAG GTCCCAAGGT CCGTCACTCA CTCCCAAGGT GAGGATATCT GTAGTTGGTC TACGAATTGA CGCAGCCGCT GGTAGTCCGG
                                Gag
-----
901  A M Q I I R D I I N E E A A D W D L Q H P Q P A P Q Q G Q L R E P S
GCATGCGAGA TTATTGCGGA CTAATCAAC GAGGAGGCTG CAGACTGGGA TTTGACGAC CCCCACCCG CCCCTCAGCA AGGGCAGCTA AGGGAGCCTT
CGCTACGCTT AATAAGCCCT GTATTAGTTG CTCTCCGAC GTCTGACCTT AAACGTCGTG GGGGTGGCG GGGGAGTCTT TCCGTCGAT TCCCTCGGAA
                                Gag
-----
1001 G S D I A C T T S S V D E Q I Q W M Y R Q Q N P I P V G N I Y R R
CCGGCAGCGA CATAGCTGGG ACTACTAGCT CCSTGGATGA ACAGATTCAA TGGATGTACA GACAGCAGAA TCCGATCCCC GTTGGCAACA TCTACCGGGG
GGCCGTGCTT GTATCGACCC TGATGATCGA GGCACCTACT TGCTAAGTT ACCTACATGT CTGTGCTCTT AGGCTAGGGG CAACCGTTGT AGATGGCCCG
                                Gag
-----
1101 W I Q L G L Q K C V R M Y N P T N I L D V K Q G P K E P F Q S Y V
CTGGATTCAA CTCGGACTTC AGAAGTGGT CAGAACTGAC AACCCACCA ATATTCTGGA TCTGAAACAG GGGCCGAAG AGCCCTTCA ATCCTACGTC
GAGCTAAGTT GAGCCTGAAG TCTCACGCA GTCCTACATG TTGGGGTGGT TATAGACCT ACACCTTTGTC CCGGCTTTTC TCGGGAAGT TAGGATGCG
                                Gag
-----
1201 D R F Y K S L R A E Q T D A A V K N W M T Q T L L I Q N A N P D C K
GACCGTTTCT ACAAAAGTCT ACGCCCGAG CAGACCGATG CCGCAGTGA GAACCTGGATG ACACAGACGC TCCTGATACA GAATGCTAAC CCTGATTGTA
CTGGCAAGA TGTTTTCAGA TSCGGGCTC GTCTGGCTAC GCGCTCACTT CTTGACCTAC TGTGTCTGCG AGGACTATGT CTTACGATTG GGACTAACAT
                                Gag
-----

```

1301 · L V L K G L G V N P T I E E M L T A C Q G V G G P G Q K A R L M A ·
 AACTCGTGCT GAAGGGCTTA GGGTAAACC CAACGCTGGA AGAARTCTTA ACCGCCTGCC AGGGAGTIGG TGGACCCGGA CAGAAGGCCG GGCTAATGGC
 TTGAGCACGA CTTCCTCAAT CCCCATTTGG GTTGGGACCT TCTTTACAAT TGCCGCACGG TCCCTCAACC ACCTGGGCTT GTCTTCGGGG CCGATTACCG
 Gag
 1401 · E A L K E A L A P V P I P F A A A Q Q R G P R K P I K C W N C G K ·
 CGAGGCGCTG AAAGAAGCAT TGGCTCCAGT ACCCATTCCT TTTGCTGCCG CACAACAGAG AGGTCCCGT AAACCGATCA AATGCTGGAA CTGPGGGGAG
 GCTCCGCGAC TTTCTTCCTA ACCGAGGTCA TGGTAAGGA AAACGACGGC GTGTTGTCTC TCCAGGGGCA TTTGGTGTAG TTAGCACTT GACACCTTC
 Gag
 1501 · E G H S A K Q C R A P R R Q G C W K C G K M D H V M A K C P D R Q A ·
 GAGGGGCACT CCCCTAACA ATGTCGAGCG CCTAGACGTC AGGGGTGTTG GAAGTCTGGT AAAATGGACC ACGTTATGSC CAATGCCCC GACAGACAAG
 CTCCCCGTA GCCGATTTGT TACAGTCCG GATCTGCG GGATCTGAG TCCCACAAAC TTTTACCTGG TGCAATACCG GTTACGGGG CTGTCGTTC
 Gag
 1601 · G F L G L G P W G K K P R N F P M A Q V H Q G L T P T A P P E D P ·
 CCGGGTTCCT CGGCTTAGGG CCTTGGGGAA AAAAGCCCAG AAACCTCCCA ATGGCGCAAG TACACCAGGG CCTGACCCCG ACCGCCCCCG CAGAGGACCC
 GGCCCAAGSA GCCCAATCCC GGAACCCCTT TTTTCGGGTC TTTGAGGGT TACCCTCTC ATGTGGTCCC GGACTGGGGC TGGCGGGGGG GTCTCCTGGG
 Gag
 1701 · A V D L L K N Y M Q L G K Q Q R E S R E K P Y K E V T E D L L H L ·
 AGCCGTAGAC CTCTTGA AAA ACTATATGCA GCTGGGGAAG CAGCAGCGCG AGAGTAGAGA GAAGCCCTAC AAGGAGGTTA CGGAGATCT GTTACACCTT
 TCGGATCTG GAGAACCTTT TGAATACGT CGACCCCTTC GTGCTGCGCG TCTCATCTCT CTTCGGGATG TTCTCCAAT GCCTTCTAGA CAATGTGGAA
 FMDV2A
 Gag Pro
 1801 · N S L F G G D Q N F D L L K L A G D V E S N P G P V L E L R Q R G P ·
 AATTCGTTAT TTGGTGTGA TCAGAAATTC GACCTGCTTA AACTTGTCTG CGACGTTGAG TCAAAATCCG GCCCTGTGCT GGAGTTGAGA CAGCGCCGGC
 TTAAGCAATA AACCACCACT AGTCTTAAAG CTGGACGAAT TTGAACGACC GCTGCAACTC AGTTTAGGCC CGGGACACCA CCTCAACTCT GTCCGCCCGC
 Pro
 1901 · Q R Q A V Q S P S E T G L L E V W Q D G P R D G Q M P R Q T G G F ·
 CCCAGCGCA CGCTGTTGAG AGCCCATCAG AGACGGGTCT ACTTGAGGTG TGGCAGGATG GCCCCCGTGA TGCACAGATG CCTCGCCAGA CGGGAGGGTT
 GGGTGGCGCT CGSACAAGTC TCGGGTAGTC TGTGCCAGA TGAACCTCCAC ACCGTCCCTAC CCGGGGCACI ACCTGTCTAC GGAGCGGTCT GCCCTCCCAA
 Pro
 2001 · F R P W S M G K E A P Q F P H G S S A S G A D A N C S P R G P S C ·
 CTTCGACCC TGGAGTATGG GAAAGGAGGC CCCGAGTTC CCTCATGGCT CTCTGCTC TGGCGCGGAT GCCAATGTA GCCCCCGAGG CCCTCTCTTC
 GRAGCTGGG ACCTCATACC CTTCCTCCG GGGGCTCAAG GAGTACCA GAAGACGGAG ACCGCGCCTA CGGTTAACAT CCGGGGCTCC GGGGAGAGCG
 Pro
 2101 · G S A K E L H A V G Q A A E R K Q R E A L Q G G D R G F A A P Q F S ·
 GGCTCAGCA AGGAGCTGCA CGCAGTGGGC CAGGCAGCAG AGCSCAAACA GCGAGAAGCA CTGCAGGGCG GTGACCGTGG TTTTCCCGCC CCACAAITCA
 CCGAGTCCGT TCCTCGACCT GCGTACCCG GTCCGTCGTC TCGCGTTTGT CGCTTCTCCT GACGTCCCG CACTGGCACC AAAACCGCGG GGTGTTAAGT
 Pro
 2201 · L W R R P V V T A H I E G Q P V E V L L D T G A D D S I V T G I E ·
 CTCTGTGGCG CCGACCTCTC GTGACTGCTC ATATCGAGGG TCAGCCCGTG GAGGTTTAC TGGACACTGG CGCAGACGAT TCTATTGTA CTGGCATTGA
 CAGACACCGG GGCTGGACAG CACTGACGAG TATAGCTCCC AGTCGGGCAC CTCCAAAATG ACCTGTGACC CGCTCTGCTA AGATAACACT GACCGTAACT
 Pro
 · L G P H Y T P K I V G G I G G F I N T K E Y K N V E I E V L G K R ·

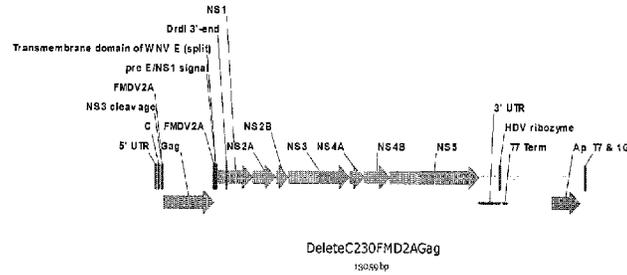
```

2301 ACTAGGCCCC CATTACACTC CMAAATCGT AGGGGGGATA GGRGGATTIA TCAACACGAA GGACTATAAG AATGTGGAGA TCGAGGTTCT CGGAAAAACC
TGATCCGGGG GTAATGTGAG GPTTTAGCA TCCCCCTAT CCTCCTAART AGTTGTGCTT CCTCATATTC TTACACCTCT AGCTCCAAGA GCCTTTGGC
FMDV2A
-----
Pro
-----
I K G T I M T G D T P I N I F G R N L L T A L G M S L N L N F D L L
2401 ATTAAGGGAA CGATTATGAC AGGCGATACA CCCATTAACA TCTTTGGACG CAATCTACTT ACGGCCCTCG GAATGAGCCT TAACCTCAAC TTCGACTTAC
TAATCCCTT GCTAATACTG TCCGCTATGT GGGTAATTGT AGAACCTGCG GTTAGATGAA TGCCGGGAGC CTACTCGGA ATTGGAGTGT AAGCTGAATG
pre E/NS1 signal
-----
FMDV2A Transmembrane domain of WNV E (split)
-----
K L A G D V E S N P G P A R D R S I A L T F L A V G G V L L F L S
2501 TCAAGCTCGC CGGAGACSTG GAGTCCAATC CCGCCCCAGC CCGGGACAGG TCCATAGCTC TCACGTTTCT CCGAGTTGGA GGAGTTCTGC TCTTCTCTC
AGTTCGAGCG GCCTCTGCAC CTCAGTTAG GGCCTGGTGC GGCCCTGTCC AGGTATCGAG AGTGCAAAGA GCGTCAACCT CCTCAAGAGC AGAAGGAGAG
Transmembrane domain of WNV E (split)
-----
NS1
-----
V N V E A D T G C A I D I S R Q E L R C G S G V F I H N D V E A W
2601 CGTGAACGTG CACGCTGACA CTGGGTGTGC CATAGACATC AGCCGGCAAG AGCTGAGATG TGGAGTGA GTGTTTATAC ACAATGATGT GGAGGCTTGG
GCACCTTSCAC GTGCGACTGT GACCCACAGC GTATCTGTAG TCGGCCGTTT TCGACTCTAC ACCTTCACCT CACAAGTATG TGTTACTACA CCTCCGAACC
NS1
-----
M D R Y K Y Y P E T P Q G L A K I I Q K A H K E G V C G L R S V S R
2701 ATGGACCGGT ACAAGTATTA CCTGAAACG CCACAAGGCC TAGCCAAGAT CATTACAGAAA GCTCATAAGG AAGGAGTGT CCGTCTACGA TCAGTTTCCA
TACCTGGCCA TGTTATAAT GGGACTTTGC GGTGTCCCG ATCGTTCTTA GTAAGTCTTT CGAGTATTC TTCTCACAC GCCAGATGCT AGTCAAGGGT
NS1
-----
L E H Q M W E A V K D E L N T L L K
2801 GACTGGAGCA TCAAATGEGG GAAGCAGTGA AGGACGAGCT GAACACTCTT TTGAAG
CTGACCTCGT AGTTTACACC CPTCGTCACT TCCTGCTCGA CTTGTGAGAA AACTTC

```

Construct 5

1. PIV-WN (Δ CprME)-SIV FMD2a Gag



2. Sequence of PIV-WN (Δ CprME)-SIV FMD2a Gag

```

-----C
                    5' UTR
-----
1  AGTAGTTCGC CTGTGTGAGC TGACAACTT AGTAGTGTIT GTGAGGATTA ACAACAATTA ACACAGTCCG AGCTGTTTCT TAGCACGAAG ATCTCGATGT
   TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCATAT TGTGTTAAT TGTGTCACGC TCGACAAAGA ATCGTGTCTTC TAGAGCTACA
-----
                                                    NS3 cleavage
   . K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R .
101 CTAAGAAACC ACGAGGGCCC GGCAAGAGCC GGGCTGTCTA TTTGCTAAAA CGCGGAATGC CCCGGGTGTT GTCCTTGATT GSACTTAAGC AAAAGAAGCG
   GATTCCTTGG TCCTCCCGGG COGTTCTCGG CCCGACAGAT AAACGATTTT GCGCCTTACG GGGCGCACAA CAGGAACTAA CCTGAATTCG TTTTCTTCGC
   FMDV2A
-----
NS3 cleavage                                     Gag
   . N F D L L K L A G D V E S N P G P M G A R N S V L S G K K A D E L .
201 AAATTTGAC CTGTTAAAC TGGCCGGGA CGTCGAAAGC AACCCCGGTC CGATGGGCGC TAGGAATAGC GTGCTTAGTG GCAAAAAGGC TGATGAACCT
   TTTAAACTG GACAATTTG ACCGCCCCCT GCAGCTTTCG TTGGGGCCAG GGTACCCGCG ATCCTTATCG CACGAATCAC CGTTTTTCCG ACTACTTGA
   Gag
-----
   E K I R L R P G G K K K Y M L K H V V W A A N E L D R F G L A E S L .
301 GAGAAGATCC GGTCCGTCG GGGCGGAAG AAGAAGTATA TGTGAAACA TGTCTGTGG GCCGCCACG AGTTAGATAG GTTTGGGCTA GCAGAGTCAT
   CTCCTTAGG CCGAGGCGG CCCGCCCTTC TTCTTCATAT ACAACTTTGT ACAGCACACC CGGCGGTTGC TCARTCTATC CAARCCCGAT CGTCTCAGTA
   Gag
-----

```

· L E N K E G C Q K I L S V L A P L V P T G S E N L K S L Y N T V C ·
 401 TGCTGABAA CAAGGAAGA TGTCAGAAGA TACTAAGTGT COTGGCACCT TGGTACCCA CGGGCTCTGA GAACTTAAAG AGTCTGTATA ACACGTGTG
 ACGAGCTTTT GTTCCTTCTT ACAGTCTTCT ATGATTACAA GGACCGTGGG AACCATGGGT GCCCCAGACT CTTGAATTC TCAGACATAT TGTGACACAC
 Gag

· V I W C I H A E E K V K H T E E A K Q I V Q R H L V V E T G T A E ·
 501 CCTGATCTGG TGCATTACAG CCGAAGAGAA AGTGAAGCAC ACCGAAGAAG CTAAGCAAAAT AGTCCAGAGA CATTGGTCTG TGGAAACCGG GACCCCGGAG
 GCAC TAGACC ACGTAAGTGC GGCCTCTCTT TCACTTCGTG TGGCTTCTTC GATTGGTTA TCACGTCTCT GTAAACCAGC ACCTTTGGCC CTGGCGGCTC
 Gag

· T M P K T S R P T A P S S G R G G N Y P V Q Q I G G N Y V H L P L S ·
 601 ACTATGCCA AACATCCCG TCCAACCGCT CCAAGTAGTG GAAGAAGAGG TAACACCCC GTTCAGCAA TCGGGGGGAA TTACGTGAT CTCCCTTTCT
 TGAACCGGT TTTGTAGGG AGTTPGGCGA GTTTCATCAC CTTCTCTCC ATTAGTGGGG CAAGTCSIT AGCCCCCTT AATGCACGTA TCGGGGAAACA
 Gag

· P R T L N A W V K L I E E K K F G A E V V P G F Q A L S E G C T P ·
 701 CACCAAGGAC CCTCAATGCA TGGGTCAAAC TCATCGAGGA AAAGAAGTTC GGAGCGGAAG TGGTCCCAGG GTTCCAGGCA CTGAGTGAAG GGTGCACCTC
 GTGTTCTCTG GGAGTTACGT ACCCAGTTTG AGTAGCTCTT TTTCTTCAAG CCTCGCCTTC ACCAGGGTCT CAAGTCCCT GACTCACTTC CCACGTGAGG
 Gag

· Y D I N Q M L N C V G D H Q A A M Q I I R D I I N E E A A D W D L ·
 801 CTATGACATC AACAGATGC TTAAC TGGCT CGGCGACCAT CAGGCGCGGA TGCAGATTAT TCGGGACATA ATCAACGAGG AGGCTGAGA CTGGGATTG
 GATCTGTAG TTAGTCTAGC AATTGACGCA GCGCTGGTA GTCCGGCGCT ACCTTAATA AGCCCTGTAT TAGTTGCTCC TCCGAGCTCI GACCCTAAC
 Gag

· Q H P Q P A P Q Q G Q L R E P S G S D I A G T T S S V D E Q I Q W M ·
 901 CAGCACCCC AACCCCGCC TCAGCAAGGG CAGCTAAGGG AGCTTCCGG CAGCGACATA GCTGGGACTA CTAGCTCCGT GGATGAACAG ATTCATGGA
 GTCCTGGGGG TTGGGGGGG AGTCGTCC C GCGATTCCC TCGGAAGGCC GTCGCTGTAT CGACCCGTAT GATCGAGGCA CCTACTTGTG TAAGTTACCT
 Gag

· Y R Q Q N P I P V G N I Y R R W I Q L G L Q K C V R M Y N P T N I ·
 1001 TGTAAGACA GCAGAATCCG ATCCCGTGG GCAACATCTA CCGCGCTGG ATTCAACTCG GACTICAGAA GTGCGTCAGA ATGTACAACC CCACCAATAT
 ACATGTCTGT CCTCTTAGCG TAGGGGCAAC CGTTGTAGAT GCGCGCGACC TAAGTTGAGC CTGAAGTCTT CACGCACTCT TACATGTTGG GGTGGTTATA
 Gag

· L D V K Q G P K E P F Q S Y V D R F Y K S L R A E Q T D A A V K N ·
 1101 TCTGGATGTG AAACAGGGGC CGAAAGAGCC CTTTCAATCC TACGTGACC GTTCTACAR AAGTCTACCG GCGAGCAGA CCGATGCGGC AGTGAAGAC
 AGACCTACAC TTTGTCCCG GCITTCCTGG GAAAGTTAG ATGCACTGGS CAAGATGTT TICAGATGCG CGGCTCGTCT GCCTACGCGC TCACCTCTTG
 Gag

· W M T Q T L L I Q N A N P D C K L V L K G L G V N P T L E E M L T A ·
 1201 TGGATGACAC AGACGCTCCT GATACAGAA GCTAACCCCT ATGTAAACT CBTGCTGAAG GGCITAGGGG TAAACCCAAC GCTGCAAGAA ATGTTAACCG
 ACCTACTGTG TCTGCGAGGA CTATGTCTTA CGATTGGGAC TAACATTTGA GCACGACTTC CCGAATCCCC ATTTGGGTTG CGACCTTCTT TACAATTGGC
 Gag

· C Q G V G G P G Q K A R L M A E A L K E A L A P V P I P F A A A Q ·
 1301 CCTGCCAGGG AGTTGGTGA CCGGACAGA AGGCCCGCT AATGGCCGAG GCGCTGAAAG AAGCATTGGC TCCACTACCC ATTCCTTTTG CTGGCGACA
 GGACGCTCC TCAACCACCT GGGCCTGTCT TCCGGGCGGA TTACCGGCTC CCGACTTTC TTCGTAACCG AGGTCATGGG TAAGGAAAC GACGGCGTGT
 Gag

· Q R G P R K P I K C W N C G K E G H S A K Q C R A P R R Q G C W K ·
 1401 ACAGAGAGT CCCCATAAC CATCAATG CTGGAAGTGT GGAAGGAGG GGCACCTCCG TAAACATGT CGAGCGCTA GAGTCAAGG GTTGTGGAG
 TGTCTCTCA GGGGCAATTG GCTAATTTAC GACCTTGACA CCCTTCTCC CCGTGGGGG ATTTGTTACA GCTCGGGAT CTGCACTCC CACAACCTC

```

                                     Gag
-----
1501  C G K M D H V M A K C P D R Q A G E L G L G P W G K K P R N F P M A
      TGTGGTAAA TGGACCAGT TPTGGCCAAA TGCCCCGACA GACAGCCGG GTTCCTCGGG TTAGGGCCTT GGGGAAAAA GCCCAGAAAC TTCCCAATGG
      ACACCATTTT ACCTGGTCCA ATACCCGGTT ACGGGGCTGT CTGTTCGGCC CAAGGAGCCC AATCCCGGAA CCCCTTTTTT CGGGTCTTTG AAGGGTTACC

                                     Gag
-----
1601  Q V H Q G L T P T A P P E D P A V D L L K N Y M Q L G K Q Q R E S
      CGCAAGTACA CCAGGGCCTG ACCCCGACCG CCCCCCAGA GGACCCAGCC GTAGACCTCT TGAAAAACTA TATGCAGCTG GGAAGCAGC AGCGCGAGAG
      GCGTTCATGT GGTCCCGGAC TGGGGCTGGC GGGGGGGTCT CCTGGGTCGG CATCTGGAGA ACTTTTTGAT ATACGTCGAC CCCTTCGTCG TCGCGCTCTC
      FMDV2A

                                     Gag
-----
1701  R E K P Y K E V T E D L L H L N S L E G G D Q N F D L L K L A G D
      TAGAGAGAAG CCTTACAAGG AGTTACGGA AGATCTGTTA CACCTTAATT CGTTATTGCG TGGTATCAG AATTCGACC TGCTTAACT TGCTGGCGAC
      ATCTCTCTTC GGGATGTTCC TCCATGCTCT TCTAGACAAT GTGGAATTA GCATTAACC ACCACTAGTC TTAAGCTGG ACGAATTTGA ACGACCGCTG
      Transmembrane domain of WNV E (split)

pre E/NS1 signal
-----
FMDV2A
-----
1801  V E S N P G P A R D R S I A L T F L A V G G V L L F L S V N V H A D
      GTTGAGTCAA ATCCGGGGCC TGCCCGGGAC AGGTCCATAG CTCTCACGTT TCCTCGCAGT GGAGGAGTTC TGCTCTTCCT CTCGGTGAAC GTGCACCGTG
      CAACTCAGTT TAGGCCCGGG ACGGGCCCTG TCCAGGTATC GAGAGTGCAA AGACCGTCAA CCTCCTCAAG ACGAGAGGA GAGGCACITG CACGTGGGAC
      NS1

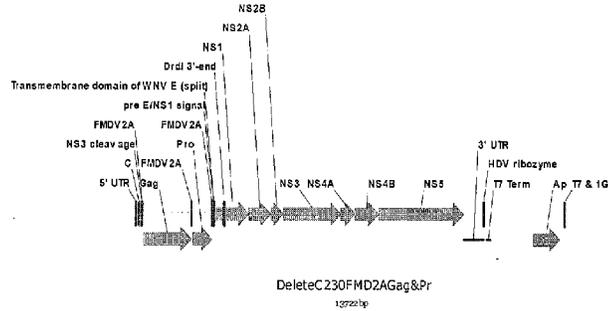
-----
1901  T G C A I D I S R Q E L R C G S G V F I H N D V E A W M D R Y K Y
      ACCTGGGTG TGCCATAGAC ATCAGCCGGC AGAGCTGAG ATGTGGAAGT SGAGTGTCA TACACAATGA TGTGGAGCT TCGATGGACC GGTACAAGTA
      TGTGACCCAC ACGGTATCTG TAGTCGGCCG TTCTCGACTC TACACCTTCA CCTCACAGT ATGTGTTACT ACACCTCGGA ACCTACCTGG CCATGTTTAT
      NS1

-----
2001  Y P E T P Q G L A K I I Q K A H K E G V C G L R S V S R L E H Q M
      TTACCTGAA ACGCCACAAG GCCTAGCCAA GATCATTGAG AAAGCTCATA AGGAAGGAGT GTCCGGTCTA CGATCAGTTT CCAGACTGGA GCATCAAATG
      AATGGGACTT TCGGTGTTT CCGATCGGTT CTAGTAAGTC TTTCGAGTAT TCCTTCTCA CACGCCAGAT GCTAGTCAAA GGTCTGACCT CGTAGTTTAC
      NS1

-----
2101  W E A V K D E L N T L L K
      TGGGAAGCAG TGAAGGACGA GCTGAACACT CTTTGAAG
      ACCCTTCGTC ACTTCCCTGCT CCACTTGTGA GAAAACTTC
    
```

Construct 6

1. PIV-WN (Δ CprME)-SIV FMD2a Gag & Pr



2. Sequence of PIV-WN (Δ CprME)-SIV fmd2A Gag & Pr (partial).

```

C
-----
5' UTR
-----
1 AGTAGTTGCG CTGTGTGAGC TGACAAACTT AGTAGTGTTT GTGAGGATTA ACAACAATTA ACACAGTGGC AGCTGTTTCT TAGCACGAAG ATCTCGATGT
  TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTGTGTTAAT TGTGTCACGC TCGACAARGA ATCGTCTCTC TAGAGCTACA
                                     NS3 cleavage
-----
C
-----
101 K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R
    CTAAGAAACC AGGAGGGCCC GGCAAGAGCC GGGCTGTCTA TTTGCTAAAA CCGGAATGC CCGCGGTGTT GTCCTTGATT GGAAGTAAAGC AAAAGAAGCG
    GATTCCTTGG TCCTCCCGGG CUGTTCCTGG CCGACAGAT AAACGATTTT GCGCCTTACG GGGCGCACAA CAGGAAGTAA CCTGAATTCG TTTTCTTCG
    FMDV2A
-----
NS3 cleavage
-----
201 N F D L L K L A G D V E S N P G P M G A R N S V L S G K K A D E L
    AAATTTGAC CTGTTAAAAC TGGCCGGGGA CGTCGAAAGC AACCCCGGTC CGATGGGCGC TAGGAATAGC GTGCTTAGTG GCAAAAAGGC TGATGAAGTT
    TTTAAAAGT CACAATTTG ACCGGCCCTT GCAGCTTTCG TTGGGGCCAG GCTACCCGCG ATCCTTATCG CACGAATCAC CGTTTTTCCG ACTACTTGAA
    Gag
-----
E K I R L R P G G K K K Y M L K H V V W A A N E L D R F G L A E S L
    
```

301 GAGAAGATCC GGCTCCGTCC GGGCGGGAAG AAGAAGTATA TGTTGAAACA TGTCGTGTGG GCGCCCAACG AGTTAGATAG GTTTGGGCTA GCAGAGTCAT
 CTCTTCTAGG CCGAGGCAGG CCCGCCCTTC TTCTTCATAT ACRACTTTGT ACAGCACACC GCGCGGTTGC TCAATCTATC CAAACCCGAT CGTCTCAGTA
 Gag

· L E N K E G C Q K I L S V L A P L V P T G S E N L K S L Y N T V C ·
 401 TGCTCGAAA CAAGGAAGGA TGTCAGAAGA TACTAAGTGT COTGGCACCT TTGGTACCCA CCGGGTCTGA GAACTTAAAG AGTCTGTATA ACACGTGTGT
 ACGAGCTTTT GTTCCTTCTT ACAGTCTTCT ATGATTACAA GGACCGTGGG AACCATGGGT GCCCCAGACT CTTGAATTC TACAGACATAT TGTGACACAC
 Gag

· V I W C I H A E E K V K H T E E A K Q I V Q R H L V V E T G T A E ·
 501 CGTGATCTGG TGCAATTCAG CCGAAGAGAA AGTGAAGCAC ACCGAAGAAG CTAAGCAAAT AGTGCAGAGA CATTGGTCTG TCGAAACCGG GACCCGCCGAG
 GCAC TAGACC ACGTAAGTGC GGCTTCTCTT TCAC TTCGTG TGGCTTCTTC GATTCCTTTA TCACGTCCTCT GTAAACCAGC ACCTTTGGCC CTGGCGGCTC
 Gag

· T M P K T S R P T A P S S G R G G N Y P V Q Q I G G N Y V H L P L S ·
 601 ACTATGCCCA AAACAATCCG TCCAACCGCT CCAAGTAGTG GAGGAGAGG TACTACCCG GTTCAGCAA TCGGCGGAA TTACGICAT CTCCTTTGT
 TGATCCGGT TTTGTAGGC AGGTTGGCGA GGTTCATCAC CTTCCTCTCC ATTGATGGG CAAGTCGTT AGCCCCCTT AATGCACGTA GAGGAAACA
 Gag

· P R T L N A W V K L I E E K K F G A E V V P G F Q A L S E G C T P ·
 701 CACCAAGGAC CCTCAATGCA TGGGTCAAAC TCATCGAGGA AAGAAGTTTC GGAGCGGAAG TGGTCCCAGG GTTCCAGGCA CTGAGTGAAG GGTGCACCTCC
 GTGGTTCCTG GGAGTTACGT ACCCAGTTTG AGTAGCTCCT TTCTTCAAG CCTCGCTTC ACCAGGTTCC CAAGGTCCGT GACTCACTTC CCACGTGAGG
 Gag

· Y D I N Q M L N C V G D H Q A A M Q I I R D I I N E E A A D W D L ·
 801 CTATGACATC AACCAGATGC TTAAC TGCCT CCGCGACCAT CAGGCCCGCA TGCAGATTAT TCGGGACATA ATCAACGAGG AAGGCTCAGA CTGGGATTTG
 GATACTGTAG TTGGTCTACG AATGACGCA GCCGCTGTA GTCCGGCTTC ACCTCTAATA AGCCCTGTAT TAGTGTCTCC TCCGAGCTCT GACCCTAARAC
 Gag

· Q H P Q P A P Q Q G Q L R E P S G S D I A G T T S S V D E Q I Q W M ·
 901 CAGCACCCCC AACCCGCCCC TCAGCAAGGG CAGCTAAGGG AGCCTTCCGG CAGCGACATA GCTGGGACTA CTAGCTCCGT GGATGAACAG ATTCAATGGA
 GTCGTGGGGG TTGGCGGGG AGTCTGTCCC GTCGATTCCC TCGGAAGGCC GTCGCTGTAT CGACCCCTGAT GATCGAGGCA CCTACTTGTG TAAGTTACCT
 Gag

· Y R Q Q N P I P V G N I Y R R W I Q L G L Q K C V R M Y N P T N I ·
 1001 TGTACAGACA GCAGAATCCG ATCCCCGTTG GCAACATCTA CCGCCGCTGG ATTCAACTCG GACTTCAGAA GTCCGTCAGA ATGTACAACC CCACCAATAT
 ACATGTCTCT CGTCTTAGGC TAGGGGCAAC CGTTGTAGAT GGCCCGGACC TTAGTTGAGC CTGAAGTCTT CACCGAGTCT TACATGTTGG GGTGGTTATA
 Gag

· L D V K Q G P K E P F Q S Y V D R F Y K S L R A E Q T D A A V K N ·
 1101 TCTGGATGTG AAACAGGGGC CGAAAGAGCC CTTTCAATCC TACGTCGACC GTTTCRCAA AAGTCTACGC GCCGAGAGA CCGATGCCGC AGTGAAGAAC
 AGACCTACAC TTTGTCCTCG GCTTTCCTCG GAAAGTTAG ATGCGCTGG CAAAGATGTT TTCAGATGCG CGCCTCTCTT GGCTACGGCG TCACITCTTG
 Gag

· W H T Q T L L I Q N A N P D C K L V L K G L G V N P T L E E M L T A ·
 1201 TGGATGACAC AGACGCTCCT GATACAGAT GCTAACCCCTG ATTGTAAACT CGTGTCAAG GCCTTAGGGG TAAACCCCAAC GCIGGAAGAA ATGTTAACC
 ACCTACTGTG TCTGCGAGGA CTATGTCTTA CGATTGGGAC TAACATITGA GCACGACTTC CCGAATCCCC ATTGGGTTG CGACCTTCTT TACAATTGGC
 Gag

· C Q G V G G P G Q K A R L M A E A L K E A L A P V P I P F A A A Q ·
 1301 CCTGCCAGGG AGTTGGTGA CCCGGACAGA AGGCCCGGCT AATGSCCGAG CGCGTAAAG AAGCATGGC TCCAGTACC ATTCCCTTTG CTGCCGCACA
 GGACGTTCC TCAACCACCT GGGCTGTCT TCCGGCCGA TTACCGGCTC CGCGACTTTC TTCGFAACCG AGGTCATGGG TAAGGAAAC GACCGCTGT
 Gag

```

1401  Q R G P R K P I K C W N C G K E G H S A K Q C R A P R R Q G C W K
ACAGAGAGGT CCCCGTAAC CGATCAAATG CTGGAACGTG GGGAGGAGG GGCACCTCGG TAAACAATST CGAGCGCCTA GCGTCAGGG GTGTTGGAG
TCTCTCTCCA GGGCAATTG GCTAGTTTAC GACCTTGACA CCTTCTCTCC CCGTGAGGGG ATTTGTTACA GCTCCCGSAT CTGAGTCCC CACAACCTTC
Gag

1501  C G K M D H V M A K C P D R Q A G F L G L G P W G K K P R N F P M A
TGTGGTAAA TGGACCAAGT TATGGCCAAA TCCCGCAGA GACAGCCGCG GTTCTCTGGG TTAGGGCCTT GSGGAAAAA GCGCAGAAC TTCCCAATGG
ACACRATTTT ACCTGGTGA AFACCGGTTT ACGGGCTGT CTGTTCGGCC CARGAGGCC AATCCGGAA CCTCTTTT CGGGTCTTG AAGGGTTACC
Gag

1601  Q V H Q G L I P T A P P E D P A V D L L K N Y M Q L G K Q Q R E S
CGCAAGTACA CCAGGGCCTG ACCCGACCG CCCCCGAGA GGACCCAGCC GTAGACTCT TGAAAACTA TATGCAGCTG GGGAGCAGC AGCCGAGAG
GCGTTCATGT GGTCCCGGAC TGGGGCTGG GGGGGGTCT CCGGGTCCG CATCTGGAGA ACTTTTGTAT ATACGTCGAC CCTTCCGTC TCGGCTCTC
FMDV2A
Gag

1701  R E K P Y K E V T E D L L H L N S L F G G D Q N F D L L K L A G D
TAGAGAGAAG CCCTACAAGG AGGTTACGGA AGATCTGTTA CACCTTAATT CGTTATTGGG TGGTATCAG AATTTGAGC TGCTAAACT TGCGGGCAG
ATCTCTCTTC GGGATGTCC TCACATGCCT TCTAGACAAT GTGCAATTA GCAATAAACC ACCACTAGTC TTAAGCTGG ACCAATTTGA ACCACCGCTG
Pro
FMDV2A

1801  V E S N P G P V L E L R Q R G P G R O A V Q S P S E T G L L E V W Q
GTTGAGTCAA ATCCGGGCC TGTCTGGAG TTGAGACAGC GCGGGCCCCA CGGGCAGGCT GTTCAGAGCC CATCAGAGAC GGTCTACTT GAGGTGIGGC
CAACTCAGTT TAGGCCCGG ACACGACCTC AACTCTGTCC GCGCCGGGTT CGCCGTCGGA CAAGTCTGG GTAGTCTCTG CCGAGATGAA CTCACACCG
Pro

1901  D G P R D G Q M P R Q T G G F F R P W S M G K E A P Q F P H G S S
AGGATGGCCC CCGTATGGA CAGATGCCTC GCCAGACGGG AGGGTCTCT CGACCTTGGG GTATGGGAAA GGAGGCCCCG CAGTTCCTCTC ATGGCTCTTC
TCCTACCGGG GGCACCTACT GTCTACGGAG CCGTCTGCCC TCCCAAGAAG GCTGGGACCT CATACCTTT CCTCCGGGGC GTCAAGGGAG TACCGAGAAG
Pro

2001  A S G A D A N C S P R G P S C G S A K E L H A V G Q A A E R K Q R
TGCCTCTGGC GCGGATGCA ATTTAGCCC CCGAGGCCCT TCTTCCGCT CAGCCAAGGA GCTGCAGCA GTGGGCCAGG CAGCAGAGCG CAACACGGA
ACGGAGACCG CGCTACGGT TAACATCGGG GGCTCCGGA AGAACGCCA GTCGGTCTCT CGACGTGGT CACCCGCTCC STCGTCTCCG GTTTGTCTCT
Pro

2101  E A L Q G G D R G F A A P Q F S L W R R P V V T A H I E G Q P V E V
GAAGCACTGC AGGGCGTGA CCGTGGTTTT GCGGCCAC AATTCATCT GTGCGCCGA CPTGCTGA CTCTCATAT CGAGGGTCAG CCCGTGGAG
CTTCTGTGAG TCCCGCACT GCCACAAA CCGCGGGGTG TTAGTCTAGA CACCSCGGT GGACAGCACT GACGATATA GCTCCAGTC GGGCACCTCC
Pro

2201  L L D T G A D D S I V T G I E L G P H Y T P K I V G G I G G F I N
TTTTACTGGA CACTGGCGA GACGATTCTA TTGTACTGG CATTGAAC TA GGGCCCAT ACACCCAAA AATCGTAGGG GGGATAGGAG GATTTATCAA
AAAATGACCT GTGACCGGT CTGCTAAGAT AACACTGACC GTAAGTTGAT CCGGGGTAA TGTGAGSTT TTAGCATCCC CCTATCTCT CTAATAGTT
Pro

2301  T K E Y K N V E I E V L G K R I K G T I M T G D T P I N I F G R N
CACGAAGGAG TATAAGAATG TGGAGATCGA GGTTCTCGGA AABCATTA AGGGAACGAT TATGACAGG GATACACCA TTAACATCTT TGGACGCAAT
GTGCTTCTC ATATCTTAC ACCTCTAGCT CCAAGACCT TTTGCGTAA TCCTTGCTA ATACTGTCG CTATCTGGT AATTGTAGAA ACCTGCGTTA
    
```

pre E/NS1 signal

FMDV2A

```

Pro
of WNV E (split)
-----
2401 I L T A L G M S L N L N F D L L K L A G D V E S N P G P A R D R S I
      C T A C T T A C G G C C C T C G G A A T G A G C C T T A A C C T C A A C T T C G A C T T A C T C A A G C T C G C C G G A G A C G T G G A G T C C A A T C C C G G C C C A G C C C G G G A C A G G T C C A
      G A T G A A T S C C G G G A G C C T T A C T C G G A A T T G C A G T T G A A G C T G A A T G A G T T C G A G C G G C C T C T G C A C C T C A G G T T A G G G C C G G G T C G G G C C C T G T C A G G T

Transmembrane domain of WNV E (split)
-----
NS1
2501 A L T F L A V G G V L L F L S V N V H A D T G C A I D I S R Q E L
      T A G C T C T C A C G I T T C T C G C A G T T G G A G G A G T T C T G C T C T T C C T C C C G T G A A C G T G C A C G C T G A C T G G G T G T G C C A T A G A C A T C A G C C G G C A G A G A G C T
      A T C G A G A G T G C A A A G A G C G T C A A C C T C C T C A A G A C C A G A A G G A G A G G C A C T T G C A C G T G C G A C T G T G A C C A C A C G G T A T C T G T A G T C G G C C G T T C T C G A
      N S 1

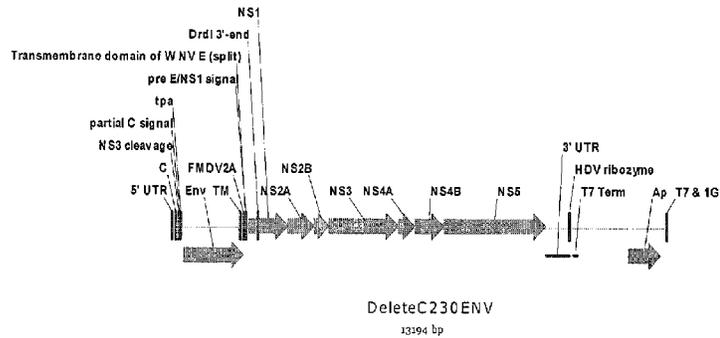
2601 R C G S G V F I H N D V E A W M D R Y K Y Y P E T P O G L A K I I
      G A G A T G T G G A A G T G G A G T G T T C A T A C A C A A T G A T G T G G A G C C T T G G A T G G A C C G G T A C A A G T A T C C C T G A A A C G C C A C A A G G C C T A G C C A A G A T C A T T
      C T C T A C A C C T T C A C C T C A C A A G T A T G T G T T A C T A C A C C T C C G A A C C T A C C T G G C C A T G T T C A T A A T G G G A C T T T G C G G T G T T C C G G A T C G G T T C T A G T A A
      N S 1

2701 Q K A H K E G V C G L R S V S R L E H Q M W E A V K D E L N T L L K
      C A G A A A G C T C A T A A G G A A G G A G T C T C C G G T C T A C A T C A G T T T C C A G A C T G G A G C A T C A A A T G T G G G A A G C A G I G A A G G A C G A G C T G A A C A C T C T T T I G A
      G T C T T C G A G T A T T C C T T C C T C A C A C C C A G A T G C T A G T C A A A G G T C T G A C C T C A G T T A C A C C C T T C G T C A C T T C C T G C T C G A C T T G T C A G A A A A C T
      N S 1

--
K
2801 A G
      T C
    
```

Construct 7

1. PIV-WN (Δ CprME)-SIV Env



2. Sequence of PIV-WN (Δ CprME)-SIV Env (partial).

```

                                                                    C
-----
                    5' UTR
-----
1  AGTAGTTCCG CTGIGTGAGC TGACAACTT AGTAGTCTTT GTGAGGATTA ACAACAATTA ACACAGTGG AGCTGTTTCT TAGCACGAAG ATCTCGATGT M S
   TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCFAAT TGTTGTTAAT TGTGTACCG TCACAAAGA ATCGTGTCTC TAGAGCTACA
-----
                                                                    NS3 cleavage
-----
· K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R ·
101 CIAAGAAAC AGGAGGSCC GGCAAGAGCC GGGCTGTCTA TTGCTAAAA CGCGGAATGC CCGCGTGTG GTCCTTGATT GGACTTAAGC AAAAGAGCG
   GATTCITTGG TCCFCCGGG CGTTCCTCG CCGACAGAT AAACGATTTT GCGCCTTACG GSGCGCRCA CAGGACTAA CCTGAATTCG TTTCTTCGC
NS3 cleavage
   ~~~~~~
   partial C signal
-----
· G G K T G I A V I M D A M K R G L C C V L L L C G A V F V T T T E
201 AGGGGGCAAG ACTGGTATAG CTGTGATCAT GGACGCCATG AAGAGGGGAC TTGTTGTGT GTCCTGTGTG TGCGGAGCTG TGTTCGTTAC AACACGGAG
   TCCCCGTTG TGACCATATC GACACTAGTA CCTGCGGTAC TTCTCCCTG AAACAACACA CGAGGACGAC ACGCCTCGAC ACAAGCAATG TTGTGCTC
-----
                    Env
    
```

```

-----
tpa
-----
301  A I Y C T Q Y V T V F Y G V P A W R N A T I P L F C A T K N R D T W
    GCGATTACT GCACCCAGTA TGTCACCGTG TTTTACGGTG TCCCGCGCTG CCGGAACGCC ACCATCCCTC TGTTTGTGC CACCAAGAAT AGAGATACGT
    CGCTAARTGA CGTGGGTCAT ACAGTGGCAC AAAATGGCCAC AGGGGGGGAC CGCTTGGCGG TGGTAGGGAG ACAAAAACAGG GTGGTTCTTA TCCTATGCA
    Env
-----
401  G T T Q C L P D N G D Y S E L A L N V T E S F D A W E N T V T E Q
    GGGGCACCAC ACAATGCCTT CCGGATAATG GCGATTACTC TGAATTAGCC CTGAACGTCA CCGAAAGTTE TGATGCTTGG GAAAATACGG TTACCGAACA
    CCCCCTGGTG TGTACGGAA GGGCTATTAC CGCTAATGAG ACTTAATCGG GACTTGCAGT GCCTTTCAAA ACIACGAACC CTTTGTATGCC AATGCTTGT
    Env
-----
501  A I E D V W Q L F E T S I K P C V K L S P L C I T M R C N K S E T
    GGGCATCGAA GATGTCTGGC AGTTATCGA AACTAGTATC AAACCTTGGC TTAAGCTGAG TCCTTTGTGC ATACAGATGC GGTGCAACAA GAGCGAANCG
    CCGGTAGCTT CTACAGACCG TCAATAAGCT TTGATCATAG TTTGGAACGC AATTGCACTC AGGAACACAG TATTGCTACG CCACGTTGTT CTCGTTTTC
    Env
-----
601  D K W G L T K S S T T T A S T T T T T A P A K I D M V N E T S S C I
    GACAAATGGG GCTTAACCAA ATCTTCAACC ACCACCCGCT CCACCACTAC SACAAACGCA CCTGCCAAGA TCGACATGGT TAACGAACCG TCATAGTTGCA
    CTGTTTACCC CGAATTGGTT TGAAGTGG TGSTGGCGGA GGTGGTGAIG GTTGTGGCGT GGACGGTCT AGCTGTACCA ATTGCTTGGC AATGCAACGT
    Env
-----
701  T H D N C T C L E Q E Q M I G C K F N M T G L K R D K T K E Y N E
    TTACCCATGA CAACTGCAGA GGCCTCGAAC AAGAACAAT GATCGGCTGT AAATTCAATA TGACCCGACT GAAGAGAGAC AAGACAAAAG AGTACAACGA
    AATGGGTACT GTTGACGTGT CCGGAGCTTG TTCTGTGTTA CTAGCCGACA TTTAAGTTAT ACTGGCTGA CTCTCTCTG TCTCTGTTTC TCATGTGCT
    Env
-----
801  T W Y S T D L V C E Q G N S T D N E S R C Y M N H C N T S I I Q E
    GACTTGCTAC AGCACCGACT TAGTGTGTGA GCAGGGGAAC TCAACCGATA ACGAGTCCCG CTGTTATATG AACCACTGCA ATACGAGCAT CATCCAAGAG
    CTGAACCATG TCGTGGCTGA ATCACACACT CGTCCCTTG AGTTGGCTAT TGCTAGGGC GACAATATAC TTGGTACCT TATGCTCGTA GTAGTTTCTC
    Env
-----
901  S C D K H Y W D T I R F R Y C A P P G Y A L L R C N D T N Y S G F M
    TCGTGGACA AACACTATTG GGACACTATC CGAATTAGGT ACTGTGCCCC GCGGGGTAT GCGCTTCTGC GTTGTATGA TACCAATTAC AGTGGGTCA
    AGCACGCTGT TTGTATAAC CCTGTGATAG GCTAATCCA TGACACGGGG CCGCCCGATA CGGGAAGACG CAACATTAAT ATGGTTAATG TCACCCAAGT
    Env
-----
1001 P K C S K V V V S S C T R M M E T Q T S T W F G F N G T R A E N R
    TGCGAAGCTG TAGCAAGTC GTGGTCTCCT CTTGTACCGG CATGATGGAG ACGCAGACTT CCACCTGGTT TGGCTTAAAC GGAACCTGAG CTGAAAACCG
    ACGGCTTAC ATCSTTTTAC CACCACAGGA GAACAATGGG GTACTACCTC TGCGTCTGAA GGTGGACCAA ACCGAAATG CCTGTAGCTC GACTTTTGGC
    Env
-----
1101 T Y I Y W H G R D N R T I I S L N K Y Y N L T M K C R R P G N K T
    GACGTATATC TACTGGCAGC GACGAGATAA CCGAACGATC ATCTCACTGA ACAAGTACTA CAATCTGACC ATGAAATGCC GCGGCCAGG CAATAAGACG
    CTGCATATAG ATGACCGTGC CTGCTCTATT GCGTTGCTAG TAGAGTACT TGTTCATGAT GTTAGACTGG TACTTTACGG CCGCGGGTCC GTTATCTGC
    Env
-----
1201 V L P V T I M S G L V F H S Q P V N E R P N Q A W C W F G G N W K D
    GIACTTCTG TCACTATTAT GAGCGGACTT GTATTTCACT CGCAGCCGGT CAATGAGCGC CCGAACCAAG CCTGGTGTG GTTTGGAGGC AACTGGAAG
    CATGAAGGAC AGTATAATA CTCGCCTGAA CATAAAGTGA GCGTCGGCA GTTACTCGCG GGCTTGGTTC GGACCAAGC CARACCTCCG TTGACCTTTC
    Env
-----

```

```

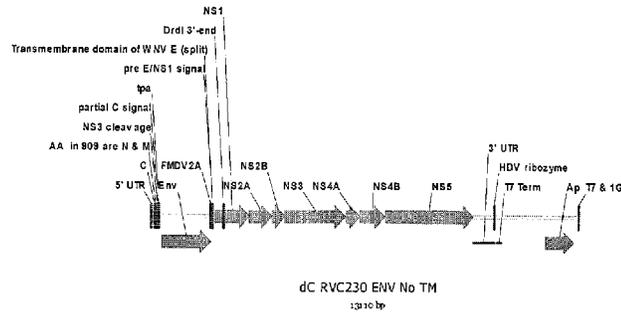
.....
1301 . A I K E V K Q T I V K H P R Y T G T N N T D K I N L T A P R G G D .
ATCGGATTAA GGAAGTAAA CAACCCATCG TARAGCATCC CCGCTACACC GGCACCAACA ATACGGATAA GATCAACCTC ACAGCCCCCTC GTGGCGGCGA
TACCCATATT CCTTCAATTT GTTTGGTAGC ATTTCCTAGG GGCAGCTGG CCGTGGTTCT TATGCCTATT CTAGTTGGAG TGTCGGGGAG CACCGCCGCT
Env
.....
1401 . P E V T F M W T N C R G E F L Y C K M N W F L N W V E D R D L T T .
TCCAGAGGTG ACCTTCATGT GGACTAACTG TCGCGGTGAA TTTCTCTACT GTAAGATGAA TTGGTTTCTG AACTGGGTGG AGGATAGGGA TCTGACAACA
AGGTCTCCAC TGAAGTACA CCGTATTGAC AGGCCACTT AAAGACATGA CATTCTACTT AACCAAGAC TTGACCCAGC TCCTATCCCT AGACTGTTGT
Env
.....
1501 Q R P K E R H R R N Y V P C H I R Q I I N T W H K V G K N V Y L P P .
CAACGGCCTA AGGAGAGGCA CCGCCGTAAC TATGTGCCTT GTCATATCAG ACAGATCATC AATACATGGC ATAAGGTGGG TAAAAAGTA TACCTCCCTC
GTTCCGGAT TCCTCTCCGT GCGCCGATTG ATACACGGAA CAGTATAGC TGCTTAGTAG TTATGTACCG TATTCCACC ATTTTGCAT ATGGAGGGAG
Env
.....
1601 . R E G D L T C N S T V T S L I A N I D W T D G N Q T N I T M S A E .
CCCGCGAGG CGACTGACA TGTAAATAGTA CAGTACCAG CCTCATGGCT AACATAGCT GGACTGATGG AAATCAGACC AACATCACA TGTACCGCA
GGCGCTCCC GCTGGACTGT ACATTATCAT GTCATTCGTC GGATAGCGA TTGTATCTGA CCGTACTACC TTTAGTCTGG TTGTAGTAT ACAGTCCGCT
Env
.....
1701 . V A E L Y R L E L G D Y K L V E I T P I G L A P T D V K R Y T T G .
GGTAGCCGAA CTGTATAGC TAGAATTCGG TGAATAAG CTCGTCGAGA TCACCCCGAT AGGGCTCGCC CCTACAGAGC TGAACGTA TACCACCGGC
CCATCGGCTT GACATATCC ATCTTGAGCC ACTGATATC GAGCAGCTCT AGTGGGGCTA TCCCGAGCGG GGATGTCTGC ACTTTGCAAT ATGGTGGCCG
Env
.....
TM
1801 G T S R N K R Y G I Y I V V G V I L L R I V I Y I V Q M L N R V R Q .
GGTACATCAA GGAACAACG CTACGGCATC TACATCGTGG TAGGGTTCAT CCTCTAAGG ATTGTCATCT ATATCGTCA GATGCTGAT AGGGTGGAGC
CCATGTAGT CCTTGTTCG GATGCCSTAG ATGTAGCACC ATCCCCAGTA GGAGATGCC TAACAGTAGA TATAGCAAGT CTACACTTA TCCCCTCCG
TM
pre E/NS1 signal
.....
Env
.....
FMDV2A Transmembrane domain of WNVE (split)
1901 . G N F D L L K L A G D V E S N P G P A R D R S I A L T F L A V G G .
AGGGCAATTT TGACCTGTA AACTGGCCG GGGAGCTGA AAGCAACCCC GGTCCGGCCC GGGACAGTC CATAGCTCTC ACGTTTCTCG CAGTTGGAGG
TCCCGTAAA ACTGGACAAT TTTGACCGC CCCTGCAGCT TTCGTTGGG CCAGGCCCGC CCCTGTCCAG GTATCGAGAG TGCAAAGAGC GTCAACCTCC
Transmembrane domain of WNVE (split)
.....
NS1
2001 . V L L F L S V N V H A D T G C A I D I S R Q E L R C G S G V F I H .
AGTTCTGCTC TTCCCTCCG TGAACGTGCA CGCTGACACT GGGGTGCCA TAGACATCAG CCGGCAAGAG CTGAGATGIG GAAGTGGAGT GTTCATACAC
TCAAGCAGAG AAGGAGAGC ACTTGCACGT CGACTGTGA CCCACGGGT ATCTGTAGTC GGCCGTTCTC GACTCTACAC CTTCACCTCA CAAGTATGTC
NS1
.....
2101 N D V E A W M D R Y K Y Y P E T P Q G L A K I I Q K A H K E G V C G .
AATGATGIG AGGCTTGGAT GGACCGGTAC AAGTATTACC CTGAACGCC ACAAGGCCA GCCAAGATCA TTCAGARAG TCATAGGAA GGAGTGTGG
TTACTACAC TCCGAACCTA CCGGCCATG TCATAAAGG GACTTTGGG TGTTCCGGAT CGGTTTAGT AAGTCTTTCG AGTATTCCTT CCTCACAGC
NS1

```

· L R S V S R L E H Q M W E A V K D E L N T L L K ·
 2201 GTCTACGATC AGTTTCCAGA CTGGAGCATC AAATGTGGGA AGCAGTGAAG GACGAGCTGA ACCTCTTTT GAAG
 CAGATGCTAG TCAAAAGTCT GACCTCGTAG TTTACACCTT TCGTCACTTC CTGCTCGACT TGTGAGAAAA CTTC

Construct 8

1. PIV-WN (ΔCprME)-SIV Env No Transmembrane (TM)



2. Sequence of PIV-WN (ΔCprME)-SIV Env No Transmembrane (partial).

C

5' UTR

1 AGTAGTTCGC CIGTGTGAGC TGACAAACTT ACTAGTGTTT GTGAGGATTA ACAACATTA ACACAGTGG AGCTGTTTCT TAGCAGGAAG ATCTCGATGT M S ·
 TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTTGTTAAT TG?GTACGC TCGACAAAGA ATCGTGCTTC TAGAGCTACA NS3 cleavage

C

· K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R ·
 101 CTAAGAAACC AGGAGGGCCC GCGAAGAGCC GSGCTGTCTA TTTCCTAAAA CGCGGATGC CCCGCGTGT GTCCTTGATT GGACTTAAGC AAAAGAAGCG
 GATTCCTTGG TCCGCCGGG CCGTTCGCG CCGACAGAT AAACGATTTT GCGCCTTACG GGGCGCACAA CAGGAATAA CCTGAATTCG TTTTCTTCGC
 NS3 cleavage
 tpa

partial C signal

.....

· G G K T G I A V I M D A M K R G L C C V L L L C G A V F V T T T E
 201 AGGGGGCAAG ACTGCIATAG CCGTGATCAT GGACGCCATG AAGAGCGGAC TTTGTTGTGT GCTCCTGCTG TCGGAGCTG TGTTCGTTAC AACCAAGGAG
 ICCTCCGTTG TGACCATATG GACACTAGTA CCGTGGGTC TCTCCCTG AAACAACACA CGAGGACGAC AGCCTCGAC ACAAGCAATG TTGTGCTCTC
 tpa

.....

Env

· A I Y C T Q Y V T V F Y G V P A W R N A T I P L F C A T K N R D T W ·
 301 GCGATTACT GCACCCASTA TGTCACCGTG TTTTACGGTG TCCCAGCGTG GCGGAACGCC ACCATCCCTC TGTTTGTGTC CACCAAGAAAT AGAGATACGT
 CGCTAAATGA CGTGGGTGAT ACAGTGGCAC AAAATGCCAC AGGGGCGGAC GCGTTGCGG TGGTAGGGAG ACAAAACACG CTGGTCTTTA TCTCTATGCA

.....

· G T T Q C L P D N G D Y S E L A L N V T E S F D A W E N T V T E Q ·
 401 GGGGACCAC ACAATGCGTT CCGGATATG GCGATTACT TGAATTAGCC CTGAACGTCA CGGAAAGTTT TGATGCTTGG GAAATACGG TTACCGAACA
 CCGGTGGTG TGTACGGAA GGGCTATTAC CCGTAATGAG ACTTAATCGG GACTTGCAGT GCCTTTCAAA ACTACGAACC CTTTATGCC AATGGCTTGT

.....

· A I E D V W Q L F E T S I K P C V K L S P L C I T M R C N K S E T ·
 501 GGCCATCGAA GATGCTGGC AGTTATTGCA AACTAGTATC AAACCTTCCG TTAAGCTGAG TCCTTGTGTC ATAACGATGC GGTGCAACA GAGCGAACG
 CCGTAGCTT CTACAGACC TCAATAAGCT TTGATCATAG TTTGGAACCG AATTCGACTC AGGAAACACG TATTGCTACG CCACGTGTGT CTCGCTTTC

.....

· D K W G L T K S S T T T A S T T T T T A P A K I D M V N E T S S C I ·
 601 GACAAATGGG GCTTACCAA ATCTTCAACC ACCACCGCCT CCACCACTAC GACACCCGA CCTGCCAAGA TCGACATGGT TAACGAAACC TCTAGTTGCA
 CTCITTACC CGAATGGTT TAGAAGTTGG TGGTGGCGA GGTGGTATG CTGTTGGCTT GGACGGTCT AGCTGTACCA ATTGCTTGG AGATCAACGT

.....

· T H D N C T G L E Q E Q M I G C K F N M T G L K R D K T K E Y N E ·
 701 TTACCCATGA CAACTGCACA GGCCTCGAAC AAGAACAAT GATCGGCTGT AAATCAATA TGACCGGACT GAAGAGAGAC AAGACAAAG AGTACRACGA
 AATGGGIACT GTTGCAGTGT CCGGAGCTIG TCTTGTGTTA CTAGCCGACA TTTAAGTTAT ACTGGCTGA CTCCTCTCTG TTCTGTTTTT CATGTTGCT

.....

· T W Y S T D L V C E Q G N S T D N E S R C Y M N H C N T S I I Q E ·
 801 GACTTGGTAC AGCACCGACT TAGTGTGTGA GCAGGGGAAC TCACCCGATA ACGACTCCCG CTGTTATATG AACCACTGCA ATACGAGCAT CATCCPAGAG
 CTGAACCAIG TCGTGGCTGA ATCACACACT CGTCCCTTTS AGTGGCTAT TGCTCAGGGC GACAATATAC TIGGTGACGT TATGCTGTA GTAGGTTCTC

.....

· S C D K H Y W D T I R F R Y C A P P C G Y A L L R C N D T N Y S G F M ·
 901 TCSTGGGACA AACACTATTG GGACACTATC CGATTAGGT ACTGTGCCCC GCGGGGCTAT GCGCTCTGTC GTTGTAAATGA TACCAATTAC AGTGGGTTC
 AGCACGCTGT TTGTGATAAC CCGTGTATAG GCTAAATCCA TGACACGGGG CCGCCCGATA CGCGAAGACG CAACATTACT ATGCTTAATG TCACCCAAGT

.....

· P K C S K V V V S S C T R M M E T Q T S T W F G F N G T R A E N R ·
 1001 TGCCGAAGTG TAGCAAAGTC GTGGTGCCT CTTGTACCG CATGATGGAG ACGGAGACTT CCACCTGTTT TGGCTTTAAC GGAAGTCCAG CTGAAAACCG
 ACGGCTTAC ATCGTTTCAG CACCACAGGA GAACATGGGC GTACTACCTC TCGCTCTGAA GGTGGACCAA ACCGAAATTG CTTTGGCTC GACTTTGGC

.....

· T Y I Y W H G R D N R T I I S L N K Y Y N L T M K C R R P G N K T ·
 1101 GACGTATATC TACTGGCAGC GACGAGATAA CCGAACGATC ATCTCACTGA ACAAGTACTA CAATCTGACC ATGAAATGCC GGGCCCGAGG CAATAAGACG
 CTGCATATAG ATGACCGTGC CTGCTCTATT GGCTTGTAG TAGAGTACT TGTTCATGAT GTTAGACTGG TACTTTACGG CCGCGGGTCC GTTATTCTGC

Env

1201 V L F V T I M S G L V F H S Q P V N E R P N Q A W C W F G G N W K D
 GIACITTCCTG TCACTATTAT GAGCGGACTT GTATITTCACT CGCAGCCCGT CAATGACCGC CCGAACCAAG CCTGGTGCTG GTTTGGAGGC AACTGGAAAG
 CATGAAGGAC AGTGATAATA CTCGCCTGAA CATAAAGTGA GCCTCGGCCA GTTACTCGCG GGCTTGGTTC GGACCACGAC CAAACCTCCG TTGACCTTTC

Env

1301 A I K E V K Q T I V K H P R Y T G T N N T D K I N L T A P R G G D
 ATCGGATTA GGAAGTTAAA CAAACCATCG TAAAGCATCC CCGCTACACC GGCACCAACA ATACGGATAA GATCAACCTC ACAGCCCTC GTGGCGGCGA
 TACGCTAAT CCTTCAATTT GTTTGGTAGC ATTTCTAGG GCGGATGTGG CCGTGGTGT TATGCCTATT CTAGTTGGAG TGTCGGGGAG CACCGCCGCT

Env

1401 P E V T F M W T N C R G E F L Y C K M N W F L N W V E D R D L T T
 TCCAGAGGTG ACCTTCATGT GGACTAACTG TCGCGGTGAA TTTCTTACT GTAAGATGAA TTGGTTCTG AACGGGTGCG AGGATAGGGA TCTGACAACA
 AGGTCTCCAC TGGAGTACA CCTGATGAC AGCCCCACTT AAGACATGA CATTCTACT AACCAAGAC TTGACCACG TCCTATCCCT AGACTGTGT

Env

1501 Q R P K E R H R R N Y V P C H I R Q I I N T W H K V G K N V Y L P P
 CAACGGCTA AGGAGAGGCA CGCCGTAAC TATGTGCTT GTCATACAG ACAGATCAT AATACATGGC ATRAGGTGG TAAAACGTA TACTCCCTC
 GTTGGCCGAT TCCTCTCCGT GCGGCAATG ATACCGGAA CAGTATGTC TGTCTAGTAG TTATGTACCG TATCCACCC ATTTTTGCAT ATGGAGGGAG

Env

1601 R E G D L T C N S T V T S L I A N I E W T D G N Q T N I T M S A E
 CCCGCGAGG CGACCTGACA TGTATAGTA CAGTAACCAG CCTCATCCCT AACATAGACT GGACTGATG AAATCAGACC AACATCATA TGTACCGCA
 GGGCCCTCC GCTGGACTGT ACATTTATCAT CTCATTGGTC GAGTATCGA TTGTATCTGA CCTGACTACC TTAGTCTGG TTGATGTGAT ACAGTCGGCT

Env

1701 V A E L Y R L E L G D Y K L V E I T P I G L A P T D V K R Y T T G
 GGTAGCCGAA CTGTATAGC TAGAATCGG TGAATAAAG CTCGTCGAGA TCACCCCGAT AGGGCTCGCC CCTACAGAGG TGAAACGTA TACCACCGGC
 CCATCGGCTT GACATATCCG ATCTTGAGCC ACTGATATTC GAGCAGCTCT AGTGGGGCTA TCCCGAGCGG GGATGTCTGC ACTTTGCAAT ATGGTGGCCG
 pre E/NS1 signal

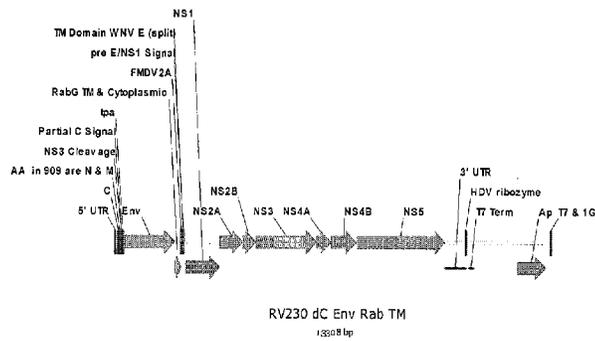
FMDV2A

| | | |
|------|--|---|
| | Env | Transmembrane domain of WNVE |
| | (split) | |
| 1801 | G T S R N K R N F D L L K L A G D V E S N P G P A R D R S I A L T F GGTACATCAA GGAACAACCG CAATTTGAC CTGTTAAAAC TCGCCGGGGA CGTCGAAAGC AACCCCGGTC CGGCCCGGGA CAGTCCATA GCTCTACGT CCATGTAGTT CCTTGTTCG GTTAAAACCTG GACAATTTTG ACCGCGCCCT GCAGCTTTCG TTGGGGCCAG GCCGGGCCCT GTCCAGGTAT CGAGAGTGCA Transmembrane domain of WNVE (split) | G T S R N K R N F D L L K L A G D V E S N P G P A R D R S I A L T F GGTACATCAA GGAACAACCG CAATTTGAC CTGTTAAAAC TCGCCGGGGA CGTCGAAAGC AACCCCGGTC CGGCCCGGGA CAGTCCATA GCTCTACGT CCATGTAGTT CCTTGTTCG GTTAAAACCTG GACAATTTTG ACCGCGCCCT GCAGCTTTCG TTGGGGCCAG GCCGGGCCCT GTCCAGGTAT CGAGAGTGCA |
| | NS1 | |
| 1901 | L A V G G V L L F L S V N V H A D T G C A I D I S R Q E L R C G S TTCTCGCAGT TGGAGGAGTT CTGCTCTCC TCCCGTGAA CGTGCAAGCT GACACTGGGT GTGCCATAGA CATCAGCCGG CAAGAGCTGA GATGTGGAAG AAGAGCCTCA ACCTCTCAA GACGAGAAGG AGAGSCACTT GCACGTGCGA CTGTGACCCA CAGCGTATCT GTAGTCGGCC GTTCTCGACT CTACACCTTC NS1 | L A V G G V L L F L S V N V H A D T G C A I D I S R Q E L R C G S TTCTCGCAGT TGGAGGAGTT CTGCTCTCC TCCCGTGAA CGTGCAAGCT GACACTGGGT GTGCCATAGA CATCAGCCGG CAAGAGCTGA GATGTGGAAG AAGAGCCTCA ACCTCTCAA GACGAGAAGG AGAGSCACTT GCACGTGCGA CTGTGACCCA CAGCGTATCT GTAGTCGGCC GTTCTCGACT CTACACCTTC NS1 |
| 2001 | G V F I H N D V E A W M D R Y K Y Y P E T P Q G L A K I I Q K A H TGGAGITTC ATACACAATG ATGTGGAGGC TTGGATGGAC CGGTACAAGT ATTACCCTGA AACGCCACAA GGCCTAGCCA AGATCATTCG GAAGCTCAT ACCPCACAAG TATGTGTAC TACACCTCCG AACCTACCTG GCCATGTCA TAATGGGACT TTGGGTGTT CCGGATCCGT TCTAGTAGT CTTCGAGTA NS1 | G V F I H N D V E A W M D R Y K Y Y P E T P Q G L A K I I Q K A H TGGAGITTC ATACACAATG ATGTGGAGGC TTGGATGGAC CGGTACAAGT ATTACCCTGA AACGCCACAA GGCCTAGCCA AGATCATTCG GAAGCTCAT ACCPCACAAG TATGTGTAC TACACCTCCG AACCTACCTG GCCATGTCA TAATGGGACT TTGGGTGTT CCGGATCCGT TCTAGTAGT CTTCGAGTA NS1 |

 K E G V C G L R S V S R L E H Q M W E A V K D E L N T L L K
 2101 AAGGAAGGAG TGTGCGGTCT ACGATCAGTT TCCAGACTGG AGCATCAAAT GTGGGAAGCA GTGAAGGACG AGCTGAACAC TCCTTTGAAG
 TTCCTTCTC ACACGGCCAGA TGCTAGTCAA AGGTCGTACC TCGTAGTTTA CACCCTTCGT CACTTCCTGC TCGACTGTG AGAAAAC TTC

Construct 9

1. PIV-WN (Δ CprME)-SIV ENV Rab G Transmembrane (TM)



2. Sequence of PIV-WN (Δ CprME)- SIV ENV Rab G Transmembrane (TM) (partial).

 5' UTR

 1 AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTTT GTGAGGATTA ACAACAATTA ACACAGTGCG AGCTGTTTCT TAGCACGAAG ATCTCGATGT
 TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTTCCTAAT TGTGTCACGC TCGACAAGA ATCGTGCTTC TAGAGCTACA
 NS3 Cleavage M S C

C

101 · K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R ·
 CTAAGAAACC AGGAGGGCCC GGCAAGAGCC GGGCTGTCTA TTGCTAANA CGCGAATGC CCCGCTGTT TCCTTGATT GGACTTAAGC AGAAAAGCG
 GATTCITTTG TCCTCCGGG CCGTTCCTGG CCCGACAGAT AARCGATTTT GCGCCTTAGC GGGCCACAA CAGGAATAA CCTGAATTCG TCTTTTTCG
 NS3 Cleavage tpa

Partial C Signal

201 · G G K T G I A V I M D A M K R G L C C V L L L C G A V F V T T T E ·
 GGGCGAAG ACAGTATTG CTGTGATCAT GGACSCCATG AAGAGGGGAC TTGTTGTGT GCTCCTGCTG TCGGAGCTG TCTTCGTTAC AACACGGGAG
 CCGCCTTTC TGCCATAAC GACACTAGTA CCGCGGATC TTCTCCCTGC AAACAACACA CGAGGACGAG ACCTCTGAC ACAAGCAAATG TTGTGCGCTC
 Env

tpa

301 · A I Y C T Q Y V T V F Y G V P A W R N A T I P L F C A T K N R D T W ·
 GCGATTACT GCACCCAGTA TGTCACCGTG TTTTACGGTG TCCCGGCTG CCGGAACGCC ACCATCCCTC TGTTTGTGC CACCAAGAAT AGAGATACGT
 CGCTAAATGA CGTGGGTCAT ACAGTGGCAC AAAATGCCAC AGGGCGGGAC CGCCTTGCGG TGGTAGGAG ACAAAACAG GTGGTCTTA TCTTATGCA
 Env

401 · G T T Q C L P D N G D Y S E L A L N V T E S F D A W E N T V T E Q ·
 GGGCACCAC ACAATGCCIT CCGATAATG GCGATTACT TGAATTAGCC CTGAACGTCA CGAAAGTTT TGATGCTGG GAAAATACGG TTACCGAACA
 CCCCCTGGTG TGTACGGAA GGGCTATTAC CGCTAATGAG ACTTAATCGG GACTTGCAGT GCCTTCAA AACTACGAACC CTTTATGCC AATGGCTTGT
 Env

501 · A I E D V W Q L F E T S I K P C V K L S P L C I T M R C N K S E T ·
 GGCCATCGAA GATGCTGGC AGTATTGCA AACTAGTATC AAACCTTGGG TTAAGCTGAG TCCTTGTGC ATAACGATGC GGTGCACAA GAGCGAAGC
 CCGTAGCTT CTACAGACCG TCAATAAGCT TTGATCATG TTGGAACGC AATTCGACTC AGGAACAGG TATTGCTAGC CCACGTTGTT CTGCGTTTC
 Env

601 · D K W G L T K S S T T T A S T T T T T A P A K I D M V N E T S S C I ·
 GACAAATGGG GCTTAACAA ATCTCAACC ACCACCGCT CCACACTAC GACAACCGA CCTGCCAAGA TCGACATGGT TAACGAACC TCTAGTTGCA
 CTGTTTACC CGAATGGIT TRGAAGTTG TGGTGGCGR GGTGTGATG CTGTTGGCT GACGCTTCT AGCTTACCA ATTGCTTGG AGATCAGCT
 Env

701 · T H D N C T G L E Q E Q M I G C K F N M T G L K R D K T K E Y N E ·
 TTACCCNTGA CAATCGACA GCCTCGAAC AAGACAAAT GATCGGCTGT AAATCAATA TGACCGACT GAAGAGAGC AAGACAAAAG AGTACAACGA
 AATGGTACT GTTACCGTGT CCGAGCCTG TTCTGTTTA CTAGCCGACA TTTAAGTTAT ACTGGCTGA CTTCTCTCG TTCTGTTTC TCATGTTGCT
 Env

801 · T W Y S T D L V C E Q G N S T D N E S R C Y M N H C N T S I I Q E ·
 GACTTGGTAC AGCACCGACT TAGTGTGTA GCAGGGGAC TCAACCGATA ACGAGTCCG CTGTTATATG AACCACTGCA ATACGAGCAT CATCCAAGAG
 CTGAACCATG TGTGCTGA ATCACACACT CCGCCCTTG AGTTGGCTAT TGCTCAGGGC GACAAATATC TTGGTGAGT TATGCTCGTA GTAGTTCTC
 Env

901 · S C D K H Y W D T I R F R Y C A P P G Y A L L R C N D T N Y S G F M ·
 TCGTGGACA AACATATTG GGACACTATC CGATTAGGT ACTGTGCCCC GCGGGGTAT GCGCTTCTGC GTTGAATGA TACCAATTAC AGTGGGTCA
 AGCACCGTGT TTGTGATAAC CCGTGTATG GCTAATCCA TGACACGGG CCGCCCGATA CGCGAAGAGC CAACATTACT ATGGTAAATG TCACCCAAGT
 Env

· P K C S K V V V S S C T R M M E T Q T S T W F G F N G T R A E N R ·

```

1001 TGCCGAAGTG TAGCAAAGTC GTGGTGTCCF CTTGTACCCG CATGATGGAG ACGCAGACTT CCACCTGGTT TGGCTTTAAC GGAAGCTCGAG CTGAAAACCG
ACGGCTTCAC ATCGTTTCAG CACCACAGGA GAACATGGGC GTACTACCIC TGGTCTGAA GGTGGACCAA ACCGAAATG CTTTGGACTC GACTTTTGGC
Env
-----
· T Y I Y W H G R D N R T I I S L N K Y Y N L T M K C R R P G N K T
1101 GACGTATATC TACTGGCAGG GACGAGATAA CCGAACGATC ATCTCACTGA ACAAGTACTA CAATCTGACC ATGAAATGCC GCGGCCAGG CAATAAGACG
CTGCATATAG ATGACCGTGC CTGCTCTATT GGCTTGCTAG TAGAGTGACT TGTTTCATGAT GTTAGACTGG TACTTTACGG CCGCGGGTCC GTTATTCTGC
Env
-----
V L P V T I M S G L V P H S Q P V N E R P N Q A W C W F G G N W K D
1201 GTACTTCTCG TCACATATTAT GAGCGCACTT GTATTTCACT CGCAGCCGGT CAATGAGCGC CCGAACCAAG CCIIGTGTCTG GTTGGGAGGC AACTGGRAAG
CATGAAGCAC AGTGATAATA CTCGCCGTAA CATAAAGTGA GCGTGGCCA GTTACTCGCG GGCTTGGTTC GGACCACGAC CAACCTCCG TTGACCTTTC
Env
-----
· A I K E V K Q T I V K H P R Y T G T N N T D K I N L T A P R G G D
1301 ATGGGATPAA GGAAGTAAA CAACCATCG TAAAGCATCC CCGCTACACC GGCACCAACA ATACGGATAA GATCAACCTC ACAGCCCTC GTGGGGGOGA
TACGCTAATT CCTCAATTT GTTGTGTAGC ATTTCCGTAGG GCGCATCTGG CCGTGGITGT TATGCCTATT CTAGTGGGAG TGTGGGGGAG CACCGCCCT
Env
-----
· P E V T F M W T N C R G E F L Y C K M N W F L N W V E D R D L T T
1401 TCCAGAGGTG ACCTTCAITG GGAATAACTG TCGCGGTGAA TTTCTGTACT GTAAGATGAA TTGGTTTCTG AACTGGGTGC AGGATAGGGA TGTGACAACA
AGGTCTCCAC TGAAGTACA CCTGATTGAC AGCGCCACTT AAAGACATGA CATTCTACTT AACCAAAGAC TTGACCCGAC TCCTATCCCT AGACTGGTGT
Env
-----
Q R P K E R H R R N Y V P C H I R Q I I N T W H K V G K N V Y L P P
1501 CAACGGCCTA AGGAGAGGCA CCGCCGTAAC TATGTGCCCT GTCATATCAG ACAGATATC AATACATGGC ATAGAGTGGG TAAAACCTA TACCTCCCTC
GTTCCCGGAT TCCTCTCCGT GGCGGCATTG ATACCCGSA CAGTATAGTC TGCTAGTAG TATGTACC TATCCACC ATTTTGCAT ATGGAGGGAG
Env
-----
· R E G D L T C N S T V T S L I A N I D W T D G N Q T N I T M S A E
1601 CCGCGGAGGG CAGCTGACA TGAATAGTA CAGTAACCAG CCTCATCGCT AACATAGACT GGACTGATGG AAATCAGACC AACATCACTA TGTACGCCA
GGCGCTCCG CTGACTGT ACATTATCAT GTCATTGGTC GGAGTAGGGA TTGTATCTGA CCTGACTACC TTTAGTCTGG TTGTAGTAT ACAGTCCGCT
Env
-----
· V A E L Y R L E L G D Y K L V E I T P I G L A P T D V K R Y T T G
1701 GGTAGCCGAA CTSTATAGCC TAGAATCGG TGAATAAAG CTGCTGAGA TCACCCCGAT AGGGCTCGCC CCTACAGAG TGAACGITA TACCACCGC
CCATCGGCTT GACATATCG ATCTTGAGCC ACTGATATTC GAGCAGCTCT AGTGGGCTA TCCGAGCGG GGATGTCTGC ACTTTGCAAT ATGGTGGCCG
Env
-----
RabG TM & Cytoplasmic
-----
G T S R N K R Y V L L S A G A L T A L M L I I F L M T C W R R V N R
1801 GGTACATCAA GGAACAACG CTACGTCCCT CTGAGTGGCG GTGCCTTAC CGCTTTGATG CTGATCATT TTCTGATGAC CTGCTGGCGG AGGGTGAATC
CCATGTAGTT CCTTGTTCG GATGCACGAG GACTCACGCC CACGGAAGT GCGAAACTAC GACTAGTAAA AAGACTACTG GACGACGCC TCCCATTAG
RabG TM & Cytoplasmic
-----
· S E P T Q H N L R G T G R E V S V T P Q S G K I I S S W E S Y K S
1901 GCTCCGAGCC GACACAGCAC AATCTCAGAG GGACAGCCCG GGAAGTAACT GTGACTCCGC AATCTGGCAA GATTATTACT AGTGGGAGA GTTACAAGTC
CGAGCTCGG CTGTGTGTG TTAGAGTCTC CCTGTCCGGC CCTCATCA CACTGAGGCG TTAGACCGTT CTANTATCA TCAACCTCT CAATGTTACG
FMDV2A
TM Domain WNV E (split)
-----
RabG TM & Cytoplasmic pre E/NS1 Signal
-----

```

· G G E T G L N F D L L K L A G D V E S N P G P A R D R S I A L T F
 2001 TGGAGGAGAG ACTGGGTTGA ATTTGATCT GCTCAAACIT GCAGGCGAIG TAGAATCAAA TCCTGGACCC GCCCGGGACA GGTCATAGC TCTCAGTTT
 ACCTCCTCTC TGACCCAACT TAAACTAGA CGAGTTTGAA CSTCCGCTAC ATCTTAGTTT AGGACCTGGG CGGGCCCTGT CCAGGTATCG AGAGTGCAAA
 NS1

TM Domain WNV E (split)

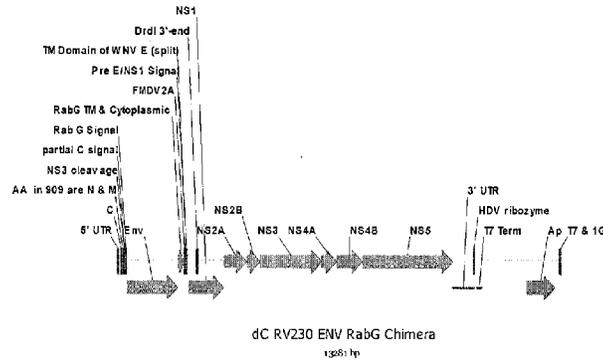
· L A V G G V L L F L S V N V H A D T G C A I D I S R Q E L R C G S G ·
 2101 CTCGCAGTTG GAGGAGTTCT GCTCTTCCTC TCCGTGAACG TGCACGCTGA CACTGSGTGT GCCATASACA TCAGCCGGCA AGAGCTGAGA TGTGGAAGTG
 GAGCCTCAAC CTCCTCAAGA CGAGAAGSAG AGGCACITGC ACGTGGCAGT CTGACCCACA CGGTATCTGT AGTCGGCCGT TCTCGACTCT ACACCTTAC
 NS1

· V F I H N D V E A W M D R Y K Y Y P E T P Q G L A K I I O K A H K ·
 2201 GAGTGTTCAT ACACAATGAT GTGGAGGCTT GGATGGACCG GTACAAGTAT TACCTGAAA CGCCACAAGG CCTAGCCAAG ATCATTGAGA AGCTCATAA
 CTCACAAGTA TGTGTTACTA CACCTCCGAA CCTACCTGGC CATGTTGATA ATGGGACTTT GCGGTGTCC GGATCGGTTT TGTAACTCT TTCAGTATT
 NS1

· E G V C G L R S V S R L E H Q M W E A V K D E L N T L L K
 2301 GGAAGGAGTG TGCGGCTTAC GATCAGTTC CAGACTGGAG CATCAATGT SGAAGCAGT GAAGGACGAG CTCARACTC TTTTGAG
 CCTTCTCAC ACGCCAGATG CTAGTCAAG STCTGACCTC GTAGTTTACA CCTTCGTC GACTTGTGAG AAAACTTC

Construct 10

1. PIV-WN (Δ CprME)-SIV Env RabG Chimera, Signal Sequence and Transmembrane (TM)



2. Sequence of PIV-WN (Δ CprME)-SIV Env RabG Chimera, Signal Sequence and Transmembrane (TM)

```

C
-----
5' UTR
-----
1 AGTAGTTGCG CTGTGTGAGC TGACAAACTT AGTAGTGTTC GTGAGGATTA ACAACAATTA ACACAGTCCG AGCTGTTTCT TAGCACGAAG ATCTCGATGT
  TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCCAAT TGITGTTAAT TGTGTACCGC TCGACAAAAG ATCCGTGCTTC TAGAGCTACA
  NS3 cleavage
  C
-----
  K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R
101 CIAAGAAACC AGGAGGGCCC GGCARGAGCC GGGCTGTCTA TTTGCTAAA CCGGGAATGC CCGCGTGTTC GTCCTTGATT GGACTTAAGC AAAAGAAGCG
  GATTCCTTGG TCCICCCGGG CCGTCTCCGG CCCGACAGAT AARCGATTTT GCGCCTTACG GGGCGCACAA CAGGAACATA CCTGAATTCG TTTTCTTCCG
  NS3 cleavage
  Rab G Signal
  partial C signal
  Env
-----
  G G K T G I A V I V P Q A L L F V P L L V F P L C F G C T Q Y V T
201 AGGGGGCAAG ACTGGTATAG CTGTGATCGT TCCTCAGGCT CTTTGTITG TACCCTTGCT GGTATTTCCC CTTTGTITG GTTGCACCCA GIATGTCACC
  TCCCCCGTTC TGACCATATC GACTAGCA AGGAGTCCGA GAAACAAAC ATGGGAACGA CCATAAAGGG GAAACGAAAC CAACGTGGGT CACACAGTGG
  Env
-----
  V F Y G V P A W R N A T I P L F C A T K N R D T W G T T Q C L P D N
  
```

301 GTGTTTACG GTGTCCTCCG CTGGCGGAAC GCCACCATCC CTCTGTTTGG TGCCACCAAG AATAGAGATA CGTGGGGCAC CACACAATGC CTCCCGGATA
CACAAAATGC CACAGGGGCG GACCGCCTTG CGTGGGTAGG GAGACAAAAC ACGGTGGTTC TTATCTCTAT GCACCCCGTG GTGTGTTACG GAAGGGCTAT
Env

401 · G D Y S E L A L N V T E S E D A W E N T V T E Q A I E D V W Q L F ·
ATGGCGATTA CTCTGAATTA GCCCTGAACG TCACGGAAAG TTTTGATGCT TGGGAAAATA CGGTTACCGA ACAGGCCATC GAAGATGTCT GGCAGTTATT
TACCGCTAAT GAGACTTAAT CGGGACTTGC AGTGCCCTTC AAAACTACGA ACCCTTTTAT GCCAATGGCT TGTCGGGTAG CTCTACAGA CCGTCAATAA
Env

501 · E T S I K P C V K L S P L C I T N R C N K S E T D K W G L T K S S ·
CGAACTAGT ATCAAACCTT GCGTAAAGT GAGTCCCTTG TGCATAACGA TGGGTGCAA CARGAGCGAA ACGGACAAAT GGGCTTAC CAAATCTTCA
GCTTTGATCA TAGTTGGAA CGCAATTCGA CTCAGGAAAC ACGTATTGCT ACGCCACGTT GTTCTCGCTT TGCCGTGTTA CCCCRAATTG GTTTAGAGT
Env

601 · I T T A S T T T T T A P A K I D M V N E T S S C I T H D N C T G L E ·
ACCACACCG CCTCACACG TACGACACCC GCACCTGCCA AGATCGCAT GGTTAACGAA ACCTCTACTT GCATTACCGA TGACAACTGC ACAGCCCTCG
TGGTGGTGGC GGAGTGGTG ATGCTGTTGG CGTGGACGCT TCTAGTCTGA CCAATTGCTT TGGAGATCAA CCTAATGGGT ACTGTTGACG TGTCCGGAGC
Env

701 · Q E Q M I G C K F N M T G L K R D K T K E Y N E T W Y S T D L V C ·
AACAAGACA AATGATCGGC TGAATTTCA ATATGACCGG ACTGAGAGA GACAAGACA AAGAGTACA CGAGACTTGG TACAGCACCG ACTTAGTGTG
TTGTTCTTGT TTACTAGCCG ACATTTAAGT TATACTGGCC TGACTTCTCT CTGTTCTGTT TTCTCATGTT GCTCTGAACC ATGTCGGTGC TGAATCACAC
Env

801 · E Q G N S T D N E S R C Y M N H C N T S I I Q E S C D K H Y W D T ·
TGAGCAGGGG AACTCAACCG ATAACGAGTC CCGCTGTTAT ATGAACCACT GCAATACGAG CATCATCCAA GAGTCGTGGG ACAAACACTA TTGGGACACT
ACTCGTCCC TTGAGTTGG TATGCTCAG GCGACAATA TACTTGTGTA CTTTATGCTC GTAGTAGGTT CTCAGCACGC TGTTTGAT AACCCTGTGA
Env

901 · I R F R Y C A P P G Y A L L R C N D T N Y S G F M P K C S K V V V S ·
ATCCGATTA GGTACTGTGC CCGCCGCGG TATGCGCTTC TGCCTGTGAA TGATACCAAT TACAGTGGGT TCATGCCGAA GTGTAGCAA GTGCTGGTGT
TAGGCTAAT CCATGACACG GGGCGCCCG ATACCGGAAG ACGCAACATT ACTATGTTA ATGTCACCCA AGTACCGCTT CACATCGTT CAGCACCA
Env

1001 · S C T R M M E T Q T S T W F G F N G T R A E N R T Y I Y W H G R D ·
CCTCTGTAC CCGCATGATG GAGACCGAGA CTTCACCTG GTTGGCTTT AACGGAACCT GAGCTGAAA CCGGACGTAT ATCTACTGGC ACGGACGAGA
GGAGAACATG GGGCTACTAC CTCTGCTCT GAAGGTGGAC CAACCGAAA TTGCTTGAG CTCGACTTTT GGCCTGCATA TAGATGACCG TCCCTGCTCT
Env

1101 · N R T I I S L N K Y Y N L T M K C R R P G N K T V L P V T I M S G ·
TAACCGAAG ATCATCTAC TGAACAAGTA CTACAATCTG ACCATGAAAT GCGCGCGCC AGGCAATAAG ACGGTACTTC CTGTCACTAT TATGAGCGGA
ATTGGCTTC TAGTAGAGTG ACTTGTTCAT GATGTTAGC TGGTACTTTA CCGCGCGGG TCCGTTAATC TGCCATGAG GACAGTATA ATACTGCTCT
Env

1201 · L V F H S Q P V N E R P N Q A W C W F G G N W K D A I K E V R Q T I ·
CTGTATTTC ACTCGAGCC GGTCAATGAG CCGCCGAACC AAGCCTGGTG CTGGTTTGA GCCAAGTGA AAGATCGGAT TRAGGAGTT AAACAACCA
GAACATAAG TGAGCGTGG CAGTTACTC GCGGGCTTG TTCCGACCAC GACCAACCT CCGTTGACCT TTCTACGCTA ATTCCTTCAA TTTGTTTGGT
Env

1301 · V K H P R Y T G T N N T D K I N L T A P R G G D P E V T F M W T N ·
TCGTAAGCA TCCCGCTAC ACGGCACCA ACRAACGGA TAGATCAAC CTCACAGCC CTCGTGGCGG CGATCCAGAG GTCACTTCA TGTGGACTAA
AGCATTTCGT ASGGCGGATG TGGCCGTGT TGTATGCTT ATCTASTTG GAGTGTGGG GAGCACCGCC GCTAGGTCTC CACTGGAAGT ACACCTGATT
Env

```

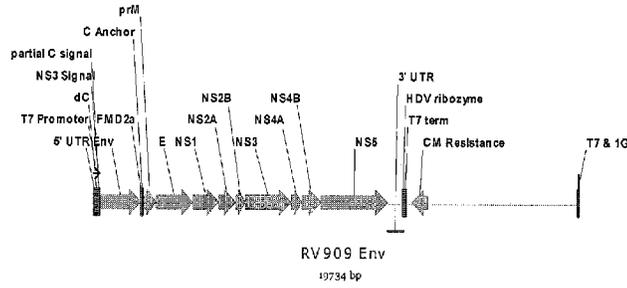
.....
* C R G E F L Y C K M N W F L N W V E D R D L T T Q R P K E R H R R
1401 CTGTCCGGT GAATTTCTGT ACTGTAAGAT GAATTSGTTT CTGAAC TGGG TCGAGGATAG GGATCTGACA ACACACACGGC CTAAGGAGAG GCACCCCGCT
GACAGGGCCA CTTAAGACA TGACATTCTA CTTAACCAAA GACTTGACCC AGCTCCTATC CCTAGACTGT TGTGTTGCGG GATTCTCTCC CGTGGCGGCA
Env
.....
* N Y V F C H I R Q I L N T W H K V G K N V Y L P P R E G D L T C N S
1501 AACTAATGC CTGTGCATAT CAGACAGATC ATCAATACAT GGCATAAGGT GGTA AAAAC GTATACCTCC CTCCCCCGCA GGGCGACCTG ACATGTAATA
TTGATACAG GAACAGTATA GTCTGTCTAG TAGTTATGTA CCGTATPCCA CCCATTTTTC CATATGGAGG GAGGGGCGCT CCCGCTGGAC TGTACATTAT
Env
.....
* T V T S L I A N I D W T D G N Q T N I T M S A E V A E L Y R L E L
1601 GTACAGTAC CAGCCTATC GCTAACATAC ACTGGACTGA TGGAAATCAG ACCAACATCA CTATGTCAGC GGAGGTAGCC GAAGCTGTATA GGCTAGAACT
CATGTCTTG GTCCGAGTAG CGATTGTATC TGACCTGACT ACCTTTAGTC TGTGTGTAAT GATACAGTCC GCTCCATCGG CTTGACATAT CCGATCTTGA
RabG TM & Cytoplasmic
Env
.....
* G D Y K L V E I T P I G L A P T D V K R Y T T G G T S R N K R Y V
1701 CGGTACTAT AAGTCCGTC AGATACACCC GATAGGGCTC GCCCTACAG ACCTCAAACG TTATACCAAC GGGGTACAT CAAGGAACAA ACGTACGTCG
GCCACTGATA TCGAGCAGC TCTAGTGGGG CTATCCCGAG CGGGGAGTTC TGCACCTTGC AATATGGTGG CCGCATGTA GTTCCTTGTG TCGATGAC
RabG TM & Cytoplasmic
.....
* L L S A G A L T A L M L I I F L M T C W R R V N R S E P T Q H N L R
1801 CTCTGACTG CCGGTGCTT GACCGTTTG ATGCTGATCA TTTTCTGAT GACCTGCTGG CAGAGGGTGA ATCGCTCGA GCCGACACAG CACAATCTCA
GAGACTCAC GCCACGGAA CTGGCAAAAC TACGACTAGT AAAAAGACTA CTGGACGACC GCCTCCACT TAGCGAGGCT CGGCTGTGTC GTGTAGAGT
FMDV2A
RabG TM & Cytoplasmic
.....
* G T G R E V S V T P Q S G K I I S S W E S Y K S G G E T G L N F D
1901 GAGGACAGG CCGGGAAGTA AGTGTGACTC CGCAATCTGG CAAGATPAT AGTGTGTTGG AGAGTTCAA CTCTGGAGGA GAGACTGGST TGAATTTTGA
CTCCCTGTC GCCCTTCAT TCACACTGAG GCGTTAGACC GTTCTAATAA TCATCAACC TCTCAATGTT CAGACTCTCT CTCTGACCCA ACTTAAACT
Pre E/NS1 Signal
.....
FMDV2A TM Domain of WN V E (split)
.....
* L L K L A G D V E S N P G P A R D R S I A L T F L A V G G V L L F
2001 TCTGCTCAA CTTCAGGGC ATGTAGAATC AAATCCTGGA CCGCCCGGG ACAGTCCAT AGCTCTCAG TTCTCCCGAG TTGGAGGAT TCTGCTCTC
AGACGAGTT GAAGTCCGC TACATCTTAG TTAGGACTT GGGGGGGGCC TGTCAGGTA TCGAGGTGC AAGAGCCCTC AACCTCTCTA AGACGAGAAG
NS1
.....
TM Domain of WN V E (split)
.....
* L S V N V H A D T G C A I D I S R Q E L R C G S G V F I H N D V E A
2101 CTCTCCGTA ACGTGCACGC TGACACTGGG TGTCCATAG ACATCAGCCG GCAAGAGCTG AGATGTGAA GTGGAGTGT CATAACAAAT GATGTGGAG
GAGAGCACT TGCACGTGGC ACTGTGACCC ACACGATATC TGTAGTGGC CGTCTCGAC TCTACACTT CACCTCACAA GTATGTGTTA CTACACTCC
NS1
.....
* W M D R Y K Y Y P E T P Q G L A K I I Q K A H K E G V C G L R S V
2201 CTTCGATGA CCGTACAAG TATTACCTG AAACGCCACA AGGCCTAGCC AAGATCATC AGAAAGCTCA TAAGGAAGGA GTCTGGCGCT TACGATCAGT
GAACCTACT GGCATGTTT AATAATGGAC TTTGGGTGT TCCGATCGG TTCTAGTAAG TCTTCGAGT ATTCTCTCT CACACGGCAG ATGCTAGTCA
NS1
.....

```

· S R L E H Q M W E A V K D E L N T L L K E N G V D L S V V V E K Q
2301 TTCCAGACTG GAGCATCAA TGTGGGAAGC AGTGAAGGAC GAGCTGAACA CTCTTTGAA GGAGAATGGT GTGGACCTTA GTGTCGTGGT TGAGAAACAA
AAGGTCGAC CTGCTAGTIT ACACCCCTCG TCACITCCTG CTCGACTTGT GAGAAAACCT CCTCTTACCA CACCTGGAAT CACAGCACCA ACTCTTTGTI

Construct 11

1. PIV-WN (Δ C)-SIV Env



2. Sequence of PIV-WN (Δ C)-SIV Env (partial).

```

                                                                    dC
                                                                    ~~~~~
                                                                    5' UTR
-----
1  S S S P V * A D K L S S V C E D * Q Q L T Q C E L F L S T K I S M S .
   AGTAGTTCCG CTGTGTGAGC TGACAAACTT ACTAGTGTTC GTGAGGATTA ACAACAATTA ACACAGTCCG AGCTGTTTCT TAGCACGAGC ATCTCGATGT
   TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTGTTAAT TGTGTACGC TCGACAAAGA ATCGTGCTTC TAGAGCTACA
                                                                    NS3 Signal
-----
                                                                    dC
-----
101  . K K P G G P G K S R A V N M L K R G M P R V L S L I G L K Q K K R .
     CTAAGAAACC AGGAGGGCCC GSCNAGGCC GGCCTGTCAA TATGCTRAAA CGCGGAATSC CCGCGTGTG GTCCCTTGATT GGACTTAAGC AAAAGAAGCG
     GATTCCTTGG TCCTCCCGGG CCGTTCCTGG CCCGACAGTT ATACGATTTT GCGCCTTACG GGGCGCACAA CAGGACTAA COTGAATTCG TTTTCTTCGC
     partial C signal
-----
NS3 Signal
-----
Env
-----
201  . G G K T G I A V I M D A M K R G L C C V L L L C G A V F V T T T E
     AGGGGGCAAG ACTGSTATAG CTGTGATCAT GGACGCCATG AAGAGGGGAC TTTGTTGTGT GCTCCTGCTG TCGCGAGCTG TGTTCGTTAC AACAAACGGAG
     TCCCGCGTTC TGACCATATC GACACTAGTA CCTGCGGTAC TTCTCCCTCG AAACAACACA CGAGGACGAC ACGCCTCGAC ACAAGCAATG TTGTGCCTC
     Env
-----

```

```

.....
A I Y C T Q Y V T V F Y G V P A W R N A T I P L F C A T K N R D T W
301 GCGATTACT GCACCCAGTA TGTCACCGTG TTTTACGGTG TCCCGGCGTG GGGGAACGCC ACCATCCCTC TGTTTTGTGC CACCAAGAAT AGAGTACGT
CGCTAAATGA CGTGGGTCAT ACAGTGGCAC AAAATGCCAC AGGGGGCGGC CCGCTTGGCG TGGTAGGGAG ACAAAACACG GTGGTTCTTA TCTCTATGCA
Env
. G T T Q C L P D N G D Y S E L A L N V T E S F D A W E N T V T E Q .
401 GGGGCACCAC ACAATGCCTT CCCGATAATG GCGATTACTC TGAATTAGCC CTGAACGTCA CGGAAAGTTT TGATGCTTGG GAAAATACGG TTACCGAACA
CCCGTGGTG TGTACGGAA GGGCTATTAC CGCTAATGAG ACTTAATCGG GACTTGCAGT GCCTTTCAAA ACTACGAACC CTTTATGCC AATGGCTTGT
Env
. A I E D V W Q L F E T S I K P C V K L S P L C I T M R C N K S E T
501 GGCCATCGAA GATGTCGGC AGTTATCGA AACTAGTATC AAACCTTGGC TTAAGCTGAG TCCTTTGTGC ATAACGATGC GGTGCAACA GAGCGAACG
CCGGTAGCTT CTACAGACCG TCAATAAGCT TTGATCATAG TTTGGAACGC AATTCGACTC AGGAAACACG TATTGCTAGC CCAGCTTGT CTGCTTTC
Env
D K W G L F K S S T T T A S T T T T T T A P A K I D M V N E T S S C I
601 GACAAATGGG GCTTAACCAA ATCTTCAACC ACCACCGCCT CCAGCACIAC GACAAACGCA CTTGCCAAGB TCGACATGGT TAACGAARCC TCTAGTTGCA
CTGTTTACC CGAATGGTT TAGAAGTGG TGSTGGCGGA GGTGTGTATG CTGTTGGCGT GGACGGTICT AGCTGTACCA ATTCCTTTGG AGATCAACCT
Env
. T H D N C T G L E Q E Q M I G C K F N M T G L K R D K T K E Y N E
701 TTACCCATGA CAACCTGACA GGCCTCGAAC AAGAACAAT GATCGGCTGT AAATTCATA TGACCGGACT GAAGAGAGAC AAGACAARAAG AGTACACGA
AATGGGTACT GTTACGCTT CCGGAGCTTG TTCTTGTTA CTAGCCGACA TTTAAGITAT ACTGGCCTGA CTTCTCTCTG TTCTGTTTC TCAITTTGCT
Env
. T W Y S T D L V C E Q G N S T D N E S R C Y M N H C N T S I I Q E
801 GACTTGGTAC AGCACCGACT TAGTGTGTGA GCAGGGGAAC TCAACCGATA ACGAGTCCCG CTGTTATATG AACCACTGCA ATACGAGCAT CATCCAAGAG
CTGAACCATG TCGTGGCTGA ATCACACACT CGTCCCTTGG AGTTGGCTAT TGCTCAGGGC GACAATATAC TTGGTGACGT TATGCTCGTA GTAGGTTCTC
Env
S C D K H Y W D T I R F R Y C A P P G Y A L L R C N D T N Y S G F M
901 TCGTGGACA AACACTATTG GGACACTATC CGATTAGGT ACTGTGCCCC GCGGGGCTAT GCGCTTCTGC GTTGTAAATGA TACCAATTAC AGTGGGTCA
AGCACGCTGT TTGTGATAAC CCTGTGATAG GCTAAATCCA TGACACGGGG GCGCCGATA CCGGAAGACG CAACATTACT ATGGTTAATG TCACCCAAGT
Env
. P K C S K V V V S S C T R M M E T Q T S T W F G F N G T R A E N R
1001 TGCCGAAGTG TAGCAAACTC GTGGTGTCTT CTGTACCCG CATGATGGAG ACGCAGACTT CCACCTGGT TGGCTTTAAC GGAACCTGAG CTGAAAACCG
ACGGCTTAC ATCGTTTCAG CACCACAGGA GAACATGGGC GTACTACCTC TGCCTCTGAA GGTGGROCAA ACCGAAATG CCTTGGCTC GACTTTTGGC
Env
. T Y I Y W H G R D N R T I I S L N K Y Y N L T M K C R R P G N K T
1101 GACGATATC TACTGGCAGC CACGAGATAA CCGAACGATC ATCTCAGTGA ACAAGTACTA CABTCTGACC ATGAAATGCC GCGCCCCAGG CAATAAGACG
CTGCATATAG ATGACCGTCC CTGCTTATT GSCTTGTAG TAGAGTACTT TGTTCATGAT GTTAGACTGG TACTTTACGG CCGCGGGTCC GTTATTCTGC
Env
V L P V T I M S G L V F H S Q P V N E R P N Q A W C W F G G N W K D
1201 GTACTTCTGT TCACTATTAT GAGCGGACTT GTATTTGACT CCGACCCGCT CAATGAGCGC CCGAACCAAG CCTGGTGTCT GTTTGGAGGC AACTGGAAAG
CATGAAGGAC AGTGATAATA CTCGCTGAA CATAAAGTGA GCGTCGGCCA GTTACTCGCG GGCTTGCITC GGACCCAGAC CAAACCTCCG TTGACCTTTC
Env
. A I K E V K Q T I V K H P R Y T G G T N N T D K I N L T A P R G G D
1301 ATCCGATTAA GGAAGTTAAA CAAACCATCG TAAAGCATCC CCGCTACACC GGCACCAACA ATACGGATAA GATCAACCTC ACAGCCCTCC GTGGCGGGA

```

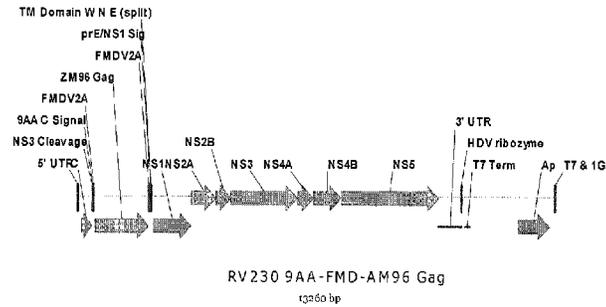
```

PACGCTAATT CCTTCAATTT GTTTGGTAGC ATTTGCTAGG GCGCATGTGG CCGTGGTTGT TATGCCCIATT CTAGTTGGAG TGTCGGGGAG CACCGCCGCT
Env
-----
1401 * P E V T F M W T N C R G E F L Y C K M N W F L N W V E D R D L T T
TCCAGAGGTG ACCITTCATG GGACTAACTG TCGCGGTGAA TTTCTGTACT CTAAGATGAA TTGTTTTCTG AACTGGGTGG AGGATAGGGA TCTGACAACA
AGGTCCTCCAC TGGAAGTACA CCTGATTGAC AGCGCCACTT AAAGACATGA CATTCTACTT AACCAAAGAC TTGACCCAGC TCCTATCCCT AGACTGTGTG
Env
-----
1501 Q R P K E R H R R N Y V P C H I R Q I I N T W H K V G K N V Y L P P
CAACGGGCTA AGGAGAGGCA CCGCCGTAC TATGTGGCTT GTCATATCAG ACAGATCATC ATACATAGCC ATAAGGTGGG TAAAAACGTA TACCTCCCTC
GITGCCGGAT TCCCTCTCCG GCGCGCATTG ATACACGGAA CAGTATAGTC TGTCTAGTAG TATGTACCC TATCCACCC ATTTTTCAT ATGGAGGGAG
Env
-----
1601 * R E G D L T C N S T V T S L I A N I D W T D G N Q T N I T M S A E
CCCGGAGGG CGACCTGACA TGTAATAGTA CAGTAACCAG CCTCATCGCT AACATAGACT GGACIGATGG AAATCAGACC AACATCACTA TGTACGCGA
GGCGCTCCC GCTGGACTGT ACATTATCAT GTCATTGGTC GGAGTAGCGA TTGTATCTGA CCTGACTACC TTTAGTCTGG TTGTAGTAT ACAGTCCGCT
Env
-----
1701 * V A E L Y R L E L G D Y K L V E I T P I G L A P T D V K R Y T T G
GGTAGCCGAA CTGTATAGGC TAGAACTCGG TGACTATAAG CTCGTGAGGA TCACCCCGAT AGGGCTCGCC CCTACAGAGC TGAACCTTA TACCACCGC
CCATCGGCTT GACATATCCG ATCTTGAGCC ACTGATATTC GAGCAGCTCT AGTGGGGCTA TCCCGAGCGG GGATGICTGC ACTTTGCAAT AIGGTGGCGC
Env
-----
1801 G T S R N K R Y G I Y I V V G V I L L R I V I Y I V Q M L N R V R Q
GGTACATCAA GGAACAARCG CTACGGCATC TACATCGTGG TAGGGTTCAT CCTCTACGG ATTGICATCT ATATCGTTCA GATGCTCAAT AGGGTGGGC
CARTGTAGTT CCTGTITTC GATGCCSTAG ATGTAGCACC ATCCCCAGTA GGAGANTGCC TAACAGTAGA TATAGCAAGT CTACGACTTA TCCCACTCCG
FMD2a
-----
Env C Anchor
-----
1901 * G N F D L L K L A G D V E S N P G P G G K T G I A V M I G L I A C
AGGGCAATTT TGACCTGTTA AAAGTGGCCG GGGACGTCGA AAGCAACCCC GGTCGGGGAG GAAAGACCGG TATTGCAGTC ATGATGGCC TGATCGCCTG
TCCCCTTAAA ACTGGACAAT TTTGACCGGC CCTGTCAGCT TTCGTTGGGG CCAGGCCCTC CTTTCTGGCC ATAACGTCAG TACTAACCGG ACTAGCCGAC
C Anchor
-----
prM
-----
2001 * V G A V T L S N F Q G K V M M T V N A T D V T D V I T I P
CGTAGGAGCA GTTACCCTCT CTAACCTCCA AGGGAAGTG ATGATACCG TAAATGCTAC TGACGTCACA GATGTCATCA CGATTCCA
GCATCCTCCT CAATGGGAGA GATTGAAGGT TCCCTCCAC TACTACTGCC ATTTACGATG ACTGCAGTGT CTACAGTAGT GCTAAGGT

```

Appendix 7

1. PIV-WN (Δ prME)-HIV Gag



2. Sequence of PIV-WN (Δ prME)-HIV Gag (partial).

```

                                                                    C
-----
                    5' UTR
-----
1  AGTAGTTCCG CTGTGTGAGC TGACAAACTT ACTAGTGTTC GTGAGGATTA ACBACAATTA ACACAGTCCG AGCTGTTTCT TAGCACGAAG ATCTCCGATG  M S
   TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCATAA TGTGTGTAAT TGTGTCAAGC TCGACAAAGA ATCGTGCTTC TAGAGCTACA
   C
-----
   K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K R A M L
101 CTAAGAAACC AGGAGGGCCC GGCAAGAGCC GGCCTGTCTA TTTGCTAAAA CGCGAATGC CCGCGTGTG TGCCTTGATT GGACTTAAGA GGGCTAIGTT
   GATTCITTTGG TCCTCCGGG CCGTTCCTCG CCCGACAGAT AAACGATTTT GGCCTTACG GGGCGCACAA CAGGAACATA CCTGAATTC CCCGATACAA
   C
-----
   S L I D G K G P I R F V L A L L A F F R F T A I A P T R A V L D R
201 GAGCCTGATC GACGGCAAGG GGCAATACG ATTTGTGTTG GCCTCTTTGG CGTTCTTCAG GTTACAGCA ATTGCTCCGA CCCGAGCAGT GCTGGATCGA
   CTCGGACTAG CTGCCGTTC CCGGTATGC TAAACACAAC CCAGAGAACC GCAAGAAGTC CAAGTTCGT TAACGAGGCT GGGCTCGTCA CCACCTAGCT
   C
-----
                                                                    NS3 Cleavage
-----
W R G V N K Q T A M K H L L S F K K E L G T L T S A I N R R S S K Q
    
```

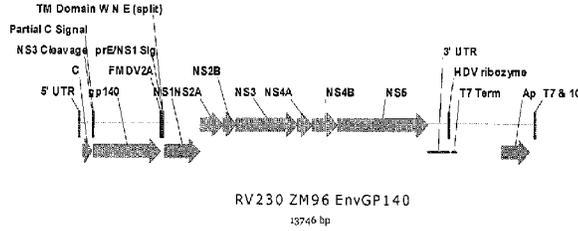
301 TGGACAGGTTG TGAACAAACA AACACGGATG AAACACCTTC TGAGTTTCAA GAAGGAACATA GGGACCTTGA CCAGTGCTAT CAATCGGCGG AGCTCAAAGC
 ACCTCTCCAC ACTTGTITTTG TTGTGCTAC TTTGTGGAAG ACTCAAAGTT CTTCCTTGAT CCTCGAACT GGTCACGATA GTTAGCCGCC TCGAGTTTCG
 9AA C Signal
 ~~~~~  
 NS3 Cleavage FMDV2A  
 ~~~~~  
 401 · K K R G G K T G I A V I N F D L L K L A G D V E S N P G P M G A R ·
 AGAAAAGCG GGGCGGAAAG ACAGGTATTG CTGTGATCAA TTTGACCTG TTAAAACCTGG CCGGGGACGT CGAAGCAAC CCCGGTCCGA TGGAGCCCG
 TCTTTTTCG CCCGCCCTTC TGCCATAAC CACTACTGTT ARAACTGGAC AATTTTGACC GGCCCTGCA GCTTTCGTTG GGGCCAGGT ACCCTCGTTC
 ZM96 Gag
 ~~~~~  
 501 · A S I L R G G K L D K W E K I R L R P G G K K R Y M I K H L V W A ·  
 ACCCAGCATC CTGAGAGGG GAAAGTGGG CAAGTGGGAG AAGATCCGGC TGAGACCTGG CGGAAGAAA CGGTACATGA TCAAGCACCT GGTGTGGGCT  
 TCGTCTGTAG GACTCTCCCG CTTCGACCT GTTACCCCTC TTCTAGGCCG ACTCTGGACC GCCTTTCITT GCCATGTACT AGTTCGTGGA CCACACCCGA  
 ZM96 Gag  
 ~~~~~  
 601 S R E L E R F A L N P G L L E T S E G C K Q I M K Q L Q P A L Q T G ·
 TCTCGGAGC TGGAAAGATT CGCCCTGAAT CCCGGCCTGC TGGAAACCAG CGAGGCTGC ARGCAATCA TGAGGAGCT GCACCTGCGC CTGCACCCG
 AGAGCCCTCG ACCTTTCATA GCGGGACTTA GGGCCGAC ACCTTGGTTC GCTCCGACG TTCTCTAGT ACTTCGTCGA CGTCGGACCG GACTCTGGC
 ZM96 Gag
 ~~~~~  
 701 · T E E L R S L Y N T V A T L Y C V H E G V E V R D T K E A L D R I ·  
 GCACCGAGGA ACTGCGGAGC CTGTACAACA CCGTGGCCAC CCTGTACTCC GTGCACGAGG GCGTGGAAAT GCGGGACACC AAAGAGGCCC TGGACCCGAT  
 CGTGGCTCT TGACGCCTCG GACTGTGTG GCCACCCGTG GGACATGACG CACGTGCTCC CGCACCTTCA CGCCCTGTGG TTCTCCGGG ACCTGGCCTA  
 ZM96 Gag  
 ~~~~~  
 801 · E E E Q N K I Q Q K I Q Q K T Q Q A A D G K V S Q N Y P I V Q N L ·
 CGAGGAAGAA CAGAACAAGA TCCAGCAGAA GATTGAGCAG AAAACCCAGC AGGCTGCCGA CGGCAAGGTG TCCAGAACT ACCCCATCGT GCAGAACCTG
 GCTCCTCTT GTCTTGTCT AGGTGCTCTT CTAAGTGCIC TTTTGGGTGC TCCGAGCGCT GCGTTCAC AGGTCCTGA TGGGTAGCA CGTCTGGAC
 ZM96 Gag
 ~~~~~  
 901 Q G Q M V H Q K L S P R T L N A W V K V I E E K A F S P E V I P M F ·  
 CAGGCCCAGA TGTGCAACA GAAGCTGTCA CCTCGGACCC TGAACCCCTG GTGAAAAGT ATCGAGGAAA AGGCCTCAG CCCTGAAGTG ATCCCATGT  
 GTCCCGTCT ACCACGTGGT CTTCGACAGT GGAGCCTGGG ACTTGGCGAC CCACCTTAC TAGCTCCTT TCCGGAATC GGGACTTAC TAGGGTACA  
 ZM96 Gag  
 ~~~~~  
 1001 · T A L S E G A T P Q D L N T M L N T V G G H Q A A M Q M L K D T I ·
 TCACAGCCCT GAGCGAGGA GCCACACCC AGGACCTGAA CACCATGCTG AACACCCTGG GAGGGCACCA GGCTGCCATG CAGATGCTGA AGGACACCAT
 APTGICGGGA CTCGCTCCT CCGTGTGGG TCTGGACTT GTGTACGAC TTGTGGCACC CTCCTGCTG CCGACGGTAC GTCTAGACT TCCTGTGGTA
 ZM96 Gag
 ~~~~~  
 1101 · N E E A A E W D R L H P V H A G P I A F G Q M R E P R G S D I A G ·  
 CAACGAAGAG GCTGCCSAGT GGGACCGCT GCACCCCTGC CATGCTGGAC CTATTGCCCC TGGCCAGATG CCGGAGCCCA GAGCCTCCA TATTGCGGC  
 GTTCTCTC CGACGGCTCA CCTGGCCGA CGTGGGACAG GTACGACCTG GATAACCGGG ACCCGTCTAC GCCCTGGGT CTCGAGGT ATACCGGCCG  
 ZM96 Gag  
 ~~~~~  
 1201 T T S T L Q E Q I A W M T S N P P I P V G D I Y K R W I I L G L N K ·
 ACCACCTCCA CACTGCAAGA ACAGATCGCC TGGATGACCA GCAACCTCC CATCCCGTG GCGACATCT ACAAGCGGTG GATCATCTG GGCCTGAACA
 TGGTGGAGT GTGACTTCT GTCTAGCGG ACCTACTGCT CSTTGGGAG GTAGGGCCAC CCGCTGTAGA TGTTCGCCAC CTAGTAGGAC CCGACTTGT
 ZM96 Gag
 ~~~~~  
 1301 · I V R M Y S P V S I L D I K Q G P K E P F R D Y V D R F E K T L R ·  
 AGATCGTGG GATGTACGC CTGTGTCCA TCTGGACAT CAAGCAGGA CCCAAAGAGC CCTTCGGGA CTACGTGGAC CGTTCCTCA AGACCTGAG

```

TCTAGCACCG CTACATGTCG GGACACAGGT AGGACCTGTA GTTCGTCCCT GGGITTTCTCG GGAAGGCCCT GATGCACCTG GCCAAGAAGT TCTGGGACTC
ZM96 Gag
-----
1401 . A E Q A T Q E V K N W M T D T L L V Q N A N P D C K T I L K A L G
AGCCGAGCAG CCCACCCAAG AGGTGAAGAA CTGGATGACC GACACCCTGC TGGTGCAGAA CGCCAACCCC GACTGCAAGA CCATCCTGAA GGCCTCGGA
TCCGCTCGTC CCGTGGGTC TCCACTTCTT GACCTACTGG CTGTCGGCAGC ACCACGCTTT GGGTTGGGG CTGACGTTCT GGTAGGACTT CCGGACCCCT
ZM96 Gag
-----
1501 P G A T L E E M M T A C Q G V G G P S H K A R V L A E A M S Q T N S
CCTGGAGCCA CCCTGGGAAG GATGATGACC GCCTGCCAGG GCCTGGGAGG ACCCAGCCAC AAGGCTCGGG TGCTGGCCGA GGCCATGAGC CAGACCAACA
GGACCTCGCT GGGACCTTCT CTACTACTGG CGGACCGTCC CGCACCTCC TGGTGGGTG TTCCGAGCCC ACGACCGCT CCGTACTCG GTCTGGTTGT
ZM96 Gag
-----
1601 . V N I L M Q K S N F K G N K R M V K C P N C G K E G H I A R N C R
GCGTGAACAT CCTGATGCAG AAGTCCAAC TCAAGGGCAA CAAGCGGAGT GTGAAGTCT TCACTGTGG AAGGAGGSC CACATTGCCA GAAACTGCAG
CGCACTTSTA GGACTACGTC TTCAGGTTGA AGTCCCGTT GTTCGCTAC CACTTCACGA AGTTGACACC TTCTCTCCG GTGTAACTGT CTTTGGCTTC
ZM96 Gag
-----
1701 . A P R K K G C W K C G K E G H Q M K D C T E R Q A N F L G K I W P
AGCCCCAAGA AAAAAGGGCT GCTGGAAGTG CGGCAAAGAG GGGCACCAGA TGAAGGACTG CACCGAGCGG CAGGCTAACT TCCTGGGCAA GATCTGGCCC
TCCGGGTTCT TTTTCCCGA CGACCTTCC GCGTTCCTC CCCGTGGTCT ACITCTCTGAC GTGGCTCGCC GTCCGATTGA AGGACCCGTT CTAGACCCGG
ZM96 Gag
-----
1801 S H K G R P G N F L Q N R P E P T A P P A E S F R F E E T T P A P K
TCCCACAAGG GCAGACCAGG CAACTTCTG CAGAACAGAC CCGAGCCAAC AGCCCTCCT GCGAGAGCT TCAGATTCA GGAARCCACC CCTGCCCCAA
AGGTTGTTC CGTCTGTCC GTTGAAGGAC GTCITGCTG GGCCTGGTG TCGGGAGGA CGGCTCTCA AGTCTAAGCT CCTTTGGTGG GACCGGGGTT
ZM96 Gag
-----
FMDV2A
-----
1901 . Q E S K D R E A L T S L K S L F G S D P L S Q N F D L L K L A G D
AGCAGGAAG CAAGGACCGG GAGGCCCTGA CCTCCCTGAA GTCCCTGTTG GCGAGCGACC CCCTGAGCCA GAATTTGAC CTGCTTAAAC TTGCTGGCGA
TCGTCCTTTC GTTCCTGGCC CTCCGGGACT GGAGGGACT CAGGGACAAG CCGTCTGTTG GCGACTCGGT CTTAAGCTG GACGAATTTG AACGACCGCT
TM Domain WN E (split)
-----
FMDV2A prE/NS1 Sig
-----
2001 . V E S N P G P A R D R S I A L T F L A V G G V L L F L S V N V H A
CGTTGAGTCA AATCCGGGGC CTGCCGGGA CAGGTCCATA GCTCTCAGT TTCTGCGAGT TGGAGGATTT CTGCTCTTCC TCTCCGTGAA CGTGCACGCT
GCAACTCAGT TTAGGCCCGG GACGGGCCCT GTCCAGGTAT CGAGAGTGA AAGAGCGTCA ACCTCCTCAA GACGAGAGG AGAGGCACCT GCACGTGCBA
NS1
-----
2101 D T G C A I D I S R Q E L R C G S G V F I H N D V E A W M D R Y K Y
GACACTGGGT GTGCCATAGA CATCAGCCGG CAAGAGCTGA GATGTGGAAG TGGAGTGTTC ATACACAATG AITGGAGGC TTGGATGGAC CGGTACAGT
CTGTGACCA CAGGTATCT GTAGTCGGCC GTTCTCGACT CTACACCTTC ACCTCACAAG TATGTGTAC TACACCTCG AACCTACCTG GCCATGTCA
NS1
-----
2201 . Y P E T P Q G L A K I I Q K A H K E G V C G L R S V S R L E H Q M
ATTACCTGA AACGCCCAA GGCCTAGCCA AGATGATTA GAAAGTCTAT AAGGAGGAG TGTCCGCTCT ACGATCAGT TCCAGACTGG AGCATCAAT
TARTGGGACT TTCCGGTGT CCGGATCGGT TCTAGTAGT CTTTCCAGTA TTCTTCTC ACACGCCAGA TGCTAGTCAA AGGTCTGACC TCCTAGTTTA
NS1
-----
W E A V K D E L N T L L K
    
```

2301 GTGGGAAGCA GTGAAGGACG AGCTGAACAC TCTTTTGAAG  
CACCCTTCGT CACTTCCATGC TCGACTTGTG AGAAAACTTC

1. PIV-WN (ΔprME)-HIV Env Gp140



2. Sequence of PIV-WN (ΔprME)-HIV Env Gp140 (partial).

```

                                     C
                                     ~~~~~
 5' UTR

1 AGTAGTTCCG CTGTGTGAGC TGACAAACTT AGTAGTGTTT GTGAGGATTA ACAACAATTA ACACAGTCGG AGCTGTTTCT TAGCACGAAG ATCTCGATGT M S
 TCATCAAGCG GACACACCCG ACTGTTTGAA TCATCACAAA CACTCCATAA TCTTGTTAAT TGTGTCCCGC TCGACAARGA ATCGTGCCTC TAGAGCTACA

 K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K R A M L
101 CTAGRAAOC AGGAGGGCCC GCAAGAGCC GGGCTGTCTA TTTGCTAAA CCGGAATGC CCCGCTGTT GTCCTTGATT GGACTIONA GGGCTATGTT
 GATTCITTTG TCCTCCCGGG CCGTTCCTGG CCCGACAGAT AACGATTTT GCGCCTTACG GGGCGACAA CAGGACTAA CCTGAATICT CCGATACAA

 S L I D G K G P I R F V L A L L A F F R F T A I A P T R A V L D R
201 GAGCCTGATC GAGGGCAGG GGC AATACG ATTGTGTGTG GCTCTCTGG CTTCTTCAG GTCACAGCA ATTCCTCCGA CCGAGCAGT GCTGGATCGA
 CTCGGACTAG CTCCCTTCC CCGTTCATGC TAACACACAC CGAGAGAACC GCAAGAAGTC CAAGTGTCTG TAACGAGGCT GGGCTCTCA CGACCTAGCT
 C

 NS3 Cleavage
                                     ~~~~~
W R G V N K Q T A M K H L L S F K K E L G T L T S A I N R R S S K Q
301 TGGAGGGTG TGAACAACA AACAGCCATG AACACCTTC TGAGTTTCAA GAAGGACTA GGGACCTTGA CCAATGCTAT CAATCGGGG AGCTCAAAGC
   ACCTCICAC ACTTGTITGT TTGTCGCTAC TTTGTGAAG ACTCAAGTIT CTTCCTTGAT CCTGGAAT GGTACAGATA GTTAGCCGCC TCGACTTTTC

```

Partial C Signal

```

NS3 Cleavage
-----
gp140
-----
401  * K K R G G K T G I A V I M G V R E I L R N W Q R W W T W G I L G F
AGAAAAAGCG GGGCGGAAAG ACAGGTATTG CTGTGATCAT GGGACTGCGG GAGATCCTGC GGAAGTGGCA GCGGTGGTGG ACCTGGGGCA TCCTGGGCTT
TCTTTTTCGC CCCGCCTTTC TGCCATAAC GACACTASTA CCCTCAGCCG CTATAGGACG CCTTGACCGT CGCCACCACC TGGACCCCGT AGGACCCGAA
gp140
-----
501  * W M L M I C N V W G N L W V T V Y Y G V P V W K E A K T T L F C A
TTGATCCTCG ATCATCTGCA ACGTGTGGGG CAACCTGTGG GTGACCGTGT ACTACGCGGT GCCCGTGTGG AAAGAGGCCA AGACCACCCT GTTCTGGGGC
AACCTACGAC TACTAGACCT TGCACACCCC GTTGGACACC CACTGACACA TGATGGCCGA CGGGACACAC TTCTCCGGT TGTGGTGGGA CAAGACGGCG
gp140
-----
601  * S D A K S Y E K E V H N V W A T H A C V P T D P N P Q E I V L G N V
AGCGACGCCA AGAGCTACGA GAAGGAAGTG CACAATGTGT GGGCCACCCA CGCCTGCGTG CCCACCGACC CCAACCCCCA GGAATCCTC CTGGGCAACG
TCGTGCGGGT TCTCGATGCT CTTCCCTCAC GTGITACACA CCCGGTGGGT GCGGACGCAC GGTGGCTGG GGTGGGGGT CCTTAGCAG GACCCGTTGC
gp140
-----
701  * T E N F N M W K N D M V D Q M H E D I I S L W D Q S L K P C V K L
TGACCGAGAA CTTCAACATG TGAAGAACG ACAITGGTGA CCAGATGCCA GAGGCATCA TCAGCCITGT GGACCAGAGC CTGAAGCCCT CGCTGAAGCT
ACTGGCTCTT GAAGTGTAC ACCTTCTTTC TGTACCACCT GGTCTACCTG CTCTGTAGT AGTCGGACAC CTTGTCTCG GACTTCGGGA CGCACTTCCA
gp140
-----
801  * T P L C V T L N C T E V N V T R N V N N S V V N N T T N V N N S M
GACCCCCCTG TGCCTGACCC TGAAGTGCAC CGAAGTGAAC GTGACCCCGA ACCTGAACAA CAGCGTGGTG AACRACCCA CCAACGTGAA TAACTCCATG
CTGGGGGAC ACGCACTGGG ACTTGACGTG GCTTCACTTG CACTGGCCCT TGCACITGTT GTCGCACCACT TTGTGTGGT GGTTCACCTT APTGAGGTAC
gp140
-----
901  * N G D M K N C S F N I T T E L K C K K K N V Y A L F Y K L D I V S L
AACGGCSACA TGAAGAATG CAGCTTCAAC ATCACCACCG AGCTGAAGGA CAAGAAAAG AACGTGTACG CCCTGTCTTA CAAGCTGGAC ATCCTGTCCC
TTGCCGCTGT ACTTCTIGAC GTCGAAGTIG TAGTGGTGGC TCGACTTCTT GTTCTTTTC TPGCACATGC GGGACAAGAT GTTCGACCTG TAGCACAGGG
gp140
-----
1001 * N E T D D S E T G N S S K Y Y R L I N C N T S A L T Q A C P K V S
TGAACGAC ACAGCAGAGC GAGACAGGCA ACAGCAGCAA GTACTACCGG CTGATCAACT GCAACACCAG CGCCCTGACC CAGGCCCTGC CCAAGGTGTC
ACTTGTCTGT TCTGTGTGCG CTCGTCCGT TGTCGTGCTT CATGATGGCC GACTAGTTGA CTTTGTGGTC CCGGACTGCG GTCCGGACGG GGTTCACAG
gp140
-----
1101 * F D P I P I H Y C A P A G Y A I L K C N N K T F N G T G P C H N V
CTTCGACCCC ATCCCATCC ACTACTGCGC CCCTCCCGCG TACGCCATCC TGAAGTGCAG CAACAAGACC TTCACGGCA CCGGCCCTG CCACAACGTC
GAAGCTGGGG TAGGGGTAGG TGATGACGCG GGGACGGCGG ATGCGGTAGG ACTTACCGTT GTTGTCTGG AAGTTGCCGT GGCCGGGGAC GGTCTTGCAC
gp140
-----
1201 * S T V Q C T H G I K P V V S T Q L L L N G S L A E E G I I I R S E N
TCCACCGTGC AGTGACCCCA CGGCATCAAG CCCGTGGTGT CCACCCAGCT GCTGCTGAAC GGCAGCCTGG CCGAGGAAGG CATCATCATC AGAAGCGAGA
AGGTGGCAGC TCACGTGGGT GCCGTAGTTC GGGCACACAA GTTGGTTCGA CGACGACTTG CCGTGGGACC GGCTCCTTCC GTAGTAGTAG TCTTCGCTCT
gp140
-----
1301 * L T N N V K T I I V H L N R S I E I V C V R P N N N T R Q S I R I
ACCTGACCAA CAACGTGAAA ACCATCATCG TGCACCTGAA CAGATCCATC GAGATCGTGT GCGTGGGGCC CAACAACAAC ACCCGGCAGA GCATCCGGAT
TGGACTGGTT GTTGACCTTT TGGTAGTAGC ACGTGGACTT GTCTAGGTAG CTCTAGCACA CGCAGCCGGG GTTGTGTGTT TGGGCCCTCT CGTAGGCCTA
gp140

```

```

.....
1401  * G P G Q T F Y A T G D I I G D I R Q A H C N I S R T N W T K T L R
      CGGCCCTGGC CAGACCTTT ACGCCACCGG CGACATCATC GGGGACATCA GACGGGCCCA CTGCACATC ACCCGGACCA ACTGGACCAA GACCCCTGGG
      GCGGGGACCG GTCTGGAAAA TCGGGTGGCC GCTGTAGTAG CCCTGTAGT CTGTCCGGGT GACCTTGTAG TCGGCCTGGT TGACCTGTTT CTGGGACGCC
      gp140
.....
1501  E V R N K L R E H F P N K N I T F K P S S G G D L E I T T H S F N C
      GAAGTGGCGA ACAAGCTGGC GGAGCACTTC CCCAACAAGA ACATCACCTT CAAGCCCGAGC TCTGGCGGCG ACCTGGAAHT CACCACCCAC AGCTTCAACT
      CTTCACGCTT GTTTCGACCG CCTGTGGAAG GGTTGTCT TGTAGTGGAA CTTCGGGTCC AGACCCGCGC TGGACCTTIA GTGGTGGGTG TCGAAGTGA
      gp140
.....
1601  * R G E F F Y C N T S G L F S I N Y T E N N T D G T P I T L P C R I
      GCAGGGGCGA GTTCTTCTAC TGCAATACCT CCGGCTGTT CAGCATCAAC TACACCGAGA ACAACACCGA CGGCACCCCC ATCACCCTGC CCGCAGAA
      CGTCCCGCT CAAGAAGATG ACCTTATGGA GCGCGGACAA GTCGTAGTGT ATGTGGCTCT TGTGTGGCT CCGCTGGGGG TAGTGGGACG GGACCTCTTA
      gp140
.....
1701  * R Q I I N M W Q E V G R A M Y A P P I E G N I A C K S D I T G L L
      CCGGCAGATC ATCAATATGT GGCAGGAGGT GGCAGGGGCC ATGTACGCC CTCCATCGA GGGCAATATC GCCTGCAAGA GCGACATCAC CGGCCTGCTG
      GCGCGTCTAG TAGTTATAGA CCGTCTCCA CCGTCCCGG TACATGCGGG GAGGGTAGCT CCCGTTATAG GGGACGTTCT CGCTGTAGTG GCGGAGGAC
      gp140
.....
1801  L V R D G G S T N D S T N N N T E I F R P A G G D M R D N W R S E L
      CTGGTGGGGG ACGGGGGCAG CACCAACGAC AGCACCAACA ACAATACCGA GATCTCCGG CCTGCCGGCG GAGACATGCG GGACAACTGG CCGAGCGAGC
      GACCACGCC TGCCCGCGTC GTGGTTGCTG TCGTGTGT TGTATGGCT CTAGAAGGCC GGAGGCGCGC CTCTGTACGC CCTGTGACC GCCTCGCTCG
      gp140
.....
1901  * Y K Y K V V E I K P L G I A P T E A K R R V V E R E K S A V G I G
      TGTACAAGTA CAAGGTGGTG GAGATCAAGC CTCTGGGCAT TGCTCCACG GAGCCCAAGC GCGGGTGGT GGAGCGGGAG AAGAGCGCGG TGGGCATCGG
      ACAITTCAT GTTCCACCAC CTCTAGTTCG GAGACCCGTA ACGAGGGTGG CTCCGGTTCG CCGCCACCA CCTCGGCCCT TCTCGGGC ACCCGTAGCC
      gp140
.....
2001  * A V F L G F L G A A C S T M G A A S I T L T A Q A R Q V L S G I V
      CGCCGTGTTT CTGGGCTTCC TGCGAGCCGC CGGAAGCACA ATGGGAGCGG CCAGCATCAC CCTGACCGCC CAGGCCCGGC AGGTGCTTC CCGCATCGTG
      GCGGCACAAA GACCCGAAG ACCCTCGGCG GCCTTCGTGT TACCGTCGG GGTGCTAGTG GGACTGGCGG GTCCGGGCGC TCCACGACAG GCGGTAGCAC
      gp140
.....
2101  Q Q Q S N L L R A I E A Q Q H L L Q L T V W G I K Q L Q T R V L A I
      CAGCAGAGA GCAACCTGCT GAGAGCCATC GAGGCTCAGC AGCACCTGCT GCAGCTGACA GTGTGGGGA TCAAGCAGCT GCAGACCCGG GTGTGGCCA
      GTCTGCTCT CGTTGGACGA CTCTCGTAG CTCCGATCG TGTGACCA CGTCGACTGT CACACCCCGT AGTTCGTGGA GCTGTGGGCC CACGACCGGT
      gp140
.....
2201  * E R Y L K D Q Q L L G L W G C S G K L I C T T A V P W N I S W S N
      TCGAGAGATA CCTGAAGGAT CACGAGTCC TGGGCTGTG GGGCTGCAG GGCAGCTGA TCTGCACCAC CGCCCTGCC TGAACATCA GCTGTCCAA
      AGCTCTCTAT GGACTTCTA GTCTCGAGG ACCCGGACAC CCCGACTCG CCGTCTGACT AGACCTGGTG CCGCACGG ACCTGTAGT CGACAGGTT
      gp140
.....
2301  * K S K T D I W D N M T W M Q W D R E I S N Y T N T I Y R L L E D S
      CAAGAGCAAG ACCGACATCT GGCACACAT GACCTGGATG CAGTGGGACC GCGAGATCAG CAACTACACC AACACATCT ACCCGCTGCT GGAAGATAGC
      GTCTCCTTC TGGCTGTAGA CCTGTGTGA CTGGACCTAC GTCACCTCG CCTCTAGT GTTGATGCG TTGTGTAGA TGGCCGACA CCTTCTATCG
      gp140
.....

```

FMDV2A

```

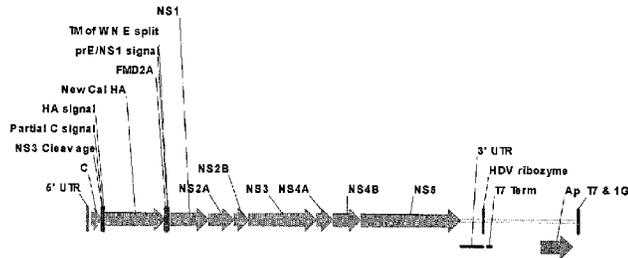
2401  Q S Q Q E Q N E K D L L A L D S W N N N F D L L K L A G D V E S N P .
      CAGAGCCAGC AGGAACAGAA CGAGAAGGAC CTGCTGGCCC TGGACAGCTG GAACAACAAT TTCGACCTGC TTAACCTTGC TGGCGACGTT GAGTCARAAC
      GTCTCGGTCC TCCTTGCTTT GCTCTTCCTG GACGACCGGG ACCTGTCCGAC CTTGTTGTTA AAGCTGGAGC AATTTGACCG ACCCGTGCAA CTCAGTTTAG
      TM Domain WN E (split)
      ~~~~~
 FMDV2A prE/NS1 sig
      ~~~~~
2501  . G P A R D R S I A L T F L A V G S V L L F L S V N V H A D T G C A .
      CCGGCCCTGC CCGGGACAGG TCCTAGCTC TCACGTTTCT CGCAGTTGGA GGAATTCCTC TCTTCCTCTC CSTGACCTGC CACGCTGACA CTGGGTGTGC
      GCCCGGAGC GGCCTCTCC AGGTATCGAG AGTGCAAAGA GCGTCAACCT CCTCAAGAGC AGAAGGAGAG GCACCTGCAC GTGGGACTGT GACCCACAGC
      NS1
      ~~~~~
2601 . I D I S R Q E L R C G S G V F I H N D V E A W M D R Y K Y Y P E T
 CATAGACATC AGCCGGCAAG AGCTGAGATG TGGAAATGGA GTGTTTATAC ACAATGATGT GGAGGCTTGG ATGGACCGGT ACAAGTATTA CCCTGAAACG
 GIATCTGTAG TCGGCCCTTC TCGACTCTAC ACCTTCACCT CACAAGTATG TGTTACTACA CCTCCGAACC TACCTGGCCA TGTTTATAAT GGGACTTTGC
 NS1
      ~~~~~
2701  P Q G L A K I I Q K A H K E G V C G L R S V S R L E H Q M W E A V K .
      CCACAAGGCC TAGCCAAGAT CATTAGAAA GCTCATAAGG AAGGAGTCTG CCGTCTACGA TCAGTTTCCA GACTGGAGCA TCAATGTGG GAAGCAGTGA
      GGTGTTCCGG ATCGGTTCTA GTAAGTCTTT CGAGTATTCC TTCCTCACAC GCCAGATGCT AGTCAAGGT CTGACCTCGT AGTTTACACC CTTCGCTACT
      NS1
      ~~~~~
2801 . D E L N T L L K
 AGGACGAGCT GAACACTCTT TTGAGC
 TCCTGCTCGA CTTGTGAGAA AACTTC

```

Appendix 8

Construct 1

1. PIV-WN ( $\Delta$ prME)-HA New Caledonia



RV230 HA New Cal Sequence  
13419bp

2. Sequence of PIV-WN ( $\Delta$ prME)-HA New Caledonia (partial).

```

C

 5' UTR

1 AGTAGTTCGC CTGTGTGAGC TGACAACCTT AGTAGTGTIT GTGAGGATTA ACAACAATTA ACACAGTGCC AGCTGTTTCT TAGCACGAAG ATCTCGATGT M S
 TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTGTGTTAAT TGTGTACAGC TCGACAAAAG ATCGTGTCTC TAGAGCTACA
 C
 K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K R A M L
101 CTAAGAAACC AGGAGGGCCC GCCAAGAGCC GGGCTGTCTA TTTGCTAAAA CGCGGAATGC CCCGCGTGT GTCCCTGATT GGACTTAAGA GAGCCATGCT
 GATTCCTTGG TCCGCCCGGG CCGTTCCTGG CCGACACAGT AAACGATTTT GCGCCTTACG GGGCGCACAA CAGGAACTAA CCTGAATTCT CTCGGTACGA
 C
 S L I D G K G P I R F V L A L L A F F R F T A I A P T R A V L D R
201 TTCGCTCATT GACGGAAGG GACCCATCCG ATTCGTACTG GCGTTGCTCG CATTCTTCCG GTTTACGGCT ATTCGCCCCA CGAGAGCGGT ACTCGACAGG
 AAGCGAGTAA CTGCCTTCC CTGGGTAGGC TAAGCATGAC CGCAACGAGC GTAAGAAGGC CAATGCGCA TAACCGCGGT GCTCTCGCCA TGAGCTGTCC
 NS3 Cleavage
 C
 W R G V N K Q T A M K H L L S F K K E L G T L T S A I N R R S S K Q
301 TGGCGCGGAG TAAACAAGCA AACAGCCATG AACACTTCT TGTCTTCAA AAAGGAAGTC GGGACCTTGA CCTCCGCCAT CAACCGAGGG AGCTCAAAAC

```

ACCGGCCTC ATTTGTTGCT TTGTCGGTAC TTGTTGAACA ACAGCAAGTT TTTCCTTGAG CCGTGGAACT GGAGGCGGTA GTTGGCTGCC TCGAGTTTTG  
 Partial C signal  
 ~~~~~  
 NS3 Cleavage  
 ~~~~~  
 HA signal  
 ~~~~~  
 401 · K K R G G K T G I A V I K A K L L V L L C T F T A T Y A D T I C I ·  
 ACAAGAAGAG GGGAGGAAAG ACGGGAATCG CAGTCAITAA GGCAAAACTG TTGGTGTTC TCTGTACATT CACTGGGAGC TACGCGGATA CAATCTGTAT  
 TCTTCTTCTC CCCTCCTTTC TGCCCTTAGC GTCAGTAATT CCGTTTTGAC AACCCACAACG AGACATGTAA GTGACGCTGC ATGCGCCTAT GTTAGACATA  
 New Cal HA  
 ~~~~~  
 501 · G Y H A N N S T D T V D T V L E K N V T V T H S V N L L E D S H N ·  
 CCGGTACCAT GCCAACAATT CGACCGACAC CGTGGATACC GTCTTGGAAA AGAATGTACC AGTGACTCAT TCGGTAACCC TCCTTGAGGA TTCGCATAAC  
 GCCCATGGTA CGGTTGTAA GCTGGCTGTG GCACCTATGG CAGCACTTTC TCTTACAGTG TCACGTAGTA AGCCATTTGG AGGAACTCCT AAGCGTATTG  
 New Cal HA  
 ~~~~~  
 601 · G K L C L L K G I A P L Q L G N C S V A G W I L G N P E C E L L I S ·  
 GGGAACTTGT GCCTTCTTAA AGGGATCGCA CCGCTGCAAC TGGGTAACCTG TTCGGTGCCT GGCTGGATTG TCGGAAACCC CGAGTGTGAA CTGCTTATCT  
 CCCTTCAACA CGGAAGAATT TCCTTAGCGT GGCGACGTTG ACCCATTTGAC AAGCCAGCGG CCGACCTAAG AGCCTTGGG GCTCACACTT GACGAATAGA  
 New Cal HA  
 ~~~~~  
 701 · K E S W S Y I V E T P N P E N G T C Y P G Y F A D Y E E L R E Q L ·  
 CAAAGGAATC GTGGTCCAT ATCGTGGAGA CACCGAACC GGAGAATGGG ACGTGTACC CTGGTTATT CCGAGACTAC GAGAACCTC GAGAACACT  
 GTTTCCTTAG CACCCAGGATA TAGCACCTCT GTGGCTTGGG CCTCTTACC TCACGATGG GACCAATAAA CGCTCTGATG CTCTTGTGAG CTCTTGTGGA  
 New Cal HA  
 ~~~~~  
 801 · S S V S S F E R F E I F P K E S S W P N H T V T G V S A S C S H N ·  
 GTCTCCGTC AGCTGGTTC AGCGATTCCA AATCTTCCG AAAGATCAT CGTGGCCGAA TCACACGGTA ACGGGTGTGT CCGCGTCACT TAGCCATAAT  
 CAGGAGGCAG TCGAGCAAGC TCGCTAAGCT TTAGAAAGGC TTTCTCAGTA GCACCCGCTT AGTGTGCCAT TGCCACACA GCGCGAGTAC ATCCGTATTA  
 New Cal HA  
 ~~~~~  
 901 · G K S S F Y R N L L W L T G K N G L Y P N L S K S Y V N N K E K E V ·  
 GGGAACTCCT CGTCTATCG CAACCTGTTG TGGCTTACTG GGAARAACGG GTTGTACCCT AATCTCAGCA AGAGCTACGT CAATAACAAA GAAAAAGAGG  
 CCCTTCAGGA GCAAGATAGC GTTGGACAAC ACCGAATGAC CCTTTTGCCT CAACATGGGA TTAGAGTCCG TCTCGATGCA GTTATTGTTT CTTTTCCTCC  
 New Cal HA  
 ~~~~~  
 1001 · L V L W G V H H P P N I G N Q R A L Y H T E N A Y V S V V S S H Y ·  
 TGCTGGTCTT GTGGGGTGTG CATCACCCAC CTAACATTGG GAATCAGAGG GCACTGTACC ACACGTAGAA TGCATACGTG AGCGTGGTGT CGAGCCACTA  
 ACGACCAGAA CACCCACAC GTAGTGGGTG GATTGTAACC CTTAGTCTCC CGTGACATGG TGTGACTCTT ACGTATGCAC TCGCACCACA GTCGGTGTAT  
 New Cal HA  
 ~~~~~  
 1101 · S R R P T P E I A K R P K V R D Q E G R I N Y Y W T L L E P G D T ·  
 TAGCCGGAGA TTCACACCAG AGATTCCGAA CCGGCCCAAA GTCCCGCACC AGGAGGGGCG GATTAACCTAC TACTGGACCC TCCTCGAGCC TGGCGATACG  
 ATCGCCCTCT AAGTGTGTC TCTAACGCTT CCGCCGGTTT CAGGCGTGG TCCTCCCGC CIAATTGATG ATGACCTGGG AGGAGCTCGG ACCGCTATGC  
 New Cal HA  
 ~~~~~  
 1201 · I I F E A N G N L I A P W Y A F A L S R G F G S G I I I S N A P M D ·  
 ATCATCTTTG AAGCGAATGG TAATCTTATC CCCCCTGGT ATGCTTTTGC GCTTTCARGA GGATTGGAT CAGGSATCAT CACATCAAAAT GCGCCGATGG  
 TAGTAGAACC TTCCCTTACC ATTAGAATAG CCGGGCACCA TACGAAACCG CGAAAGTTCT CCTAARACCTA GTCCCTAGTA GTGTAGTTTA CCGGGTACC  
 New Cal HA  
 ~~~~~  
 1301 · E C D A K C Q T P Q G A I N S S L P F Q N V H P V T I G E C P K Y ·  
 ACGAGTGCGA TGCTAACTGT CAGACTCCCC AAGCGCTAT CAATCGCTCG ITGCCCTTTC AAAACGTGCA CCCCCTAACG ATCGGAGAGT GTCCCAAGTA

TGCTCAGCCT ACGATTACACA GTCTGAGGGG TTCCGCGATA GTGAGCAGC AACGGGAAAG TTTTGCACGT GGGGCATTGC TAGCCTCTCA CAGGGTTCAT  
 New Cal HA  
 1401 · V R S A K L R M V T G L R N I P S I Q S R G L F G A I A G F I E G  
 TGTCAGATCG GCGAANCITA GGATGGTGAC CGGACTCCGC AATATCCCTT CGATCCAGTC ACGGGGATTG TTTGGAGCCA TTGCGGGCTT CATCGAAGGG  
 ACAGTCTAGC CGCTTTGAAT CCTACCACCTG GCCTGAGGCG TTATAGGGGA GCTAGGTCAG TGCCCTAAC AAACCTCGT AACGCCCGAA GTAGCTTCCC  
 New Cal HA  
 1501 G W T G M V D G W Y G Y H H Q N E Q G S G Y A A D Q K S T Q N A I N ·  
 GGCTGGACTG GAATGGTCGA TGGGTGGTAC GGTATCACC ACCAGAATGA GCAGGCTTCC GGTATGCCC CGGATCAGAA ATCGACACAG AACCAATCA  
 CCGACCTGAC CTTACCAGCT ACCCACCATG CCAATAGTGG TGGTCTTACT TGCTCCCAAG CCCATACGCC GCCTAGTCTT TAGCTGTGTC TTGCGTACT  
 New Cal HA  
 1601 · G I T N K V N S V I E K M N T Q F T A V G K E F N K L E R R M E N ·  
 ACGGGATTAC GAACAAGTA AACAGCGTCA TTGAGAAGAT GAATACACAG TTTACAGCCG TGGGGAAGA ATTCAACAAA CTCGAGGCCG GGATGGAGAA  
 TGCCCTAATG CTGTGTCCAT TTGTCCAGT AACTCTTCTA CTTATGTGTC AATGTCCGC ACCCTTTCT TAAGTTGTTT GAGCTCGGG CCTACCTCTT  
 New Cal HA  
 1701 · L N K K V D D G F L D I W T Y N A E L L V L L E N E R T L D F H D ·  
 TTTGAATAG AAAGTGGAGC ATGGTPTCCT CGATATCTGG ACGTACAATG CGGAGCTGCT TGTCCTGCTC GAAAATGAGA GGACGCTCGA CTTTCATGAC  
 AAACCTATTC TTTCACCTGC TACCAAAGGA GCTATAGACC TGCAATGTAC GCCTCGACGA ACAGGACGAG CTTTACTCTT CTTGCGGACT GAAAGTACTG  
 New Cal HA  
 1801 S N V K N L Y E K V K S Q L K N N A K E I G N G C F E F Y H K C N N ·  
 TCCAATGTA AGAACCTTTA CGAGAAGCTG AAGTCCCAAT TGAAGAATAA CGCCAAGSAA ATTGGAAACG GCTGCTTCGA ATCTACCAC AAATGCAACA  
 AGGTTACACT TCTTGGAAAT GCTCTCCAC TTCAGGGTTA ACTTCTTATT GCGGTTCCCT TAACCTTTGC CGACGAAGCT TAAGATGCTG TTTACGTTGT  
 New Cal HA  
 1901 · E C M E S V K N G T Y D Y P K Y S E E S K L N R E K I D G V K L E ·  
 ATGAGTGCAT GGAATCGGTC AAAAATGGAA CATATGATTA TCCCAATAC TCGGAGGAGT CAAAGCTTAA TAGGGAGAAA ATTGATGGGG TAAAATGTA  
 TACTACGTA CCTTAGCCAG TTTTACCTT GTATACTAAT AGGGTTTATG AGCCTCCTCA GTTTCGAATT ATCCCTCTTT TAACCTACCC ATTTTGAAT  
 New Cal HA  
 2001 · S M G V Y Q I L A I Y S T V A S S L V L L V S L G A I S F W M C S ·  
 GAGCATGGGT GTATATCAGA TCCTGGCAAT CTACTCAACC GTGGCGTGT CACTGGTACT CCTCGTTC CTGGGCGCA TTAGCTTTG CATGTGTTG  
 CTCGTACCA CATATAGTCT AGGACCGTTA GATGAGTTGG CACCCGACCA GTACCATGA GGAGCACAGG GACCCGCGCT AATCGAARAC CTACACARGC  
 New Cal HA  
 FMD2A TM of WN E split  
 2101 N G S L Q C R I C I N F D L L K L A G D V E S N F G P A R D R S I A ·  
 AATGGATCGC TCCAGTCCCG CATCTGCATC AACTTTGACC TGCTGAGCTC CGCGGGTGAC GTCGAATCCA ACCCAGGGCC AGCCCGGGC AGAAGCATTG  
 TTACCTAGCG AGGTACGGG GTAGAGTAG TTGAACCTGG ACGACTTGA GCGCCCACTG CAGCTTAGGT TGGTCCCGG TCGGGCCCTG TCTTCGTAAC  
 TM of WN E split  
 NS1  
 2201 · L T F L A V G G V L L F L S V N V H A D T G C A I D I S R Q E L R ·  
 CGCTCACTTT TCCGCGGTA GGAGGTGTGC TGTGTTCCT GTCAGTGAAC GTCCACGAG ACACGGGATG CCGGATTGAT ATCTCCAGAC AAGAATTGAG  
 GCGAGTGAAG AGAGCGCCAT CCTCCACAG ACAACAAGGA CACTCCTTG CAGGTGCTC TGTGCCCTAC GCGCTAATA TAGAGGCTG TCTTAACTC  
 NS1  
 · C G S G V F I H N D V E A W M D R Y K Y Y P E T P Q G L A K I I Q

```

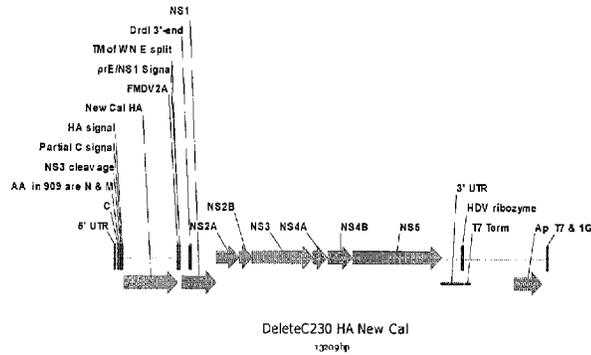
2301 GTGCGGCTCG GGGTCTTTA TCCATAACGA CGTGGAGGCG TGGATGGACA GGTATAAGTA TTACCCGTAA ACGCCGCAGG GACTTGCGAA AATCATTGAG
CACGCCCAGC CCCAGAAAT AGGTATTGCT GCACCTCCGC ACCTACCTGT CCATATTCAT AATGGGACTT TCGGGCGTCC CTGAACGCTT TTAGTAAGTC
NS1

X A H K E G V C G L R S V S R L E H Q M W E A V K D E L N T L L K
2401 AAAGCCGATA AGCAAGGTGT GTGTGGATTG AGATCAGTCT CACGCCITGA GCACCAGATG TGGACCCCTG TCAAGGATGA ATTGAACACA CTTTGAAGG
TTTCGGGTAT TCCTTCCACA CACACCTAAC TCTAGTCAGA GTGCGGAACT CGTGETCTAC ACCCTCCGAC AGTTCCTACT TAACCTTGCT GAAAACTTC

```

**Construct 2**

1. PIV-WN ( $\Delta$ CprME)-HA New Caledonia



2. Sequence of PIV-WN ( $\Delta$ CprME)-HA New Caledonia (partial).

```

C

5' UTR

1 AGTAGITCGC CTGTGTGACC TCACAAACTT AGTAGTGTTT GTGAGGATTA ACAACRATTA ACACAGTCCG AGCTGTTTCT TAGCACGAAG ATCTGGATGT M S
TCATCAAGCG GACACACTCG ACTGTTTCAA TCATCACAAA CACTCCTAAT TGTGTGTAAT TGTGTCACGC TCGACAAAGA ATCGTGTTC TAGAGTACA
NS3 cleavage

```

C

101 · K K P G G P G K S R A V Y L L K R G M P R V L S L I G L K Q K K R ·  
 CTAAGAAACC AGGAGGSCCC GGCAGAGCC GGCCTGTCTA TTGCTAAAA CCGGGAATSC CCCGCGTGT GTCCCTGATT GGACTTAAGC AGAAGAAGAG  
 GATTCTTTGG TCTCCCGGG CCGTTCTCGG CCGACACAGAT AAACGATTTT GCGCCTTACG GGGCCACAA CAGGAACATA CCTGAATTCG TCTTCTTCTC  
 Partial C signal  
 New Cal HA

NS3 cleavage HA signal

201 · G G K T G I A V I K A K L L V L L C T F T A T Y A D T I C I G Y H ·  
 GGGAGAAAG ACGGGAATCG CACTCATTAA GGCAAAACCTG TTGCTGTTCG TCTGTACATT CACTGCGAGG TACCGGGATA CAATCIGTAT CCGGTACCAT  
 CCCTCCTTC TGCCCTTAGC GTCAGTAATT CCGTTTGGAC AACACAAACG AGACATGTAA GTGACCGTGC ATGCCCTAT GTTAGACATA GCCCATGGTA  
 New Cal HA

301 · A N N S T D T V D T V L E K N V T V T H S V N L L E D S H N G K L C ·  
 GCCAACAATT CGACCGACAC CGTGGATACC GTCTGGAAA AGAATGTCCG AGTGACTCAT TCGGTAACC TCCTTGAGGA TTCCGATAC GGAAGATCT  
 CGGTTGTTAA GCTGGCIGTG GCACCTATGG CAGAACCTTT TCTTACAGTG TCACTACATA AGCCTTTGG AGGAACCTCT AAGCGTATGG CCCTTCARCA  
 New Cal HA

401 · L L K G I A P L Q L G N C S V A G N I L S N P E C E L L I S K E S ·  
 GCCTTCTTAA AGGGATCGCA CCGCTGCAAC TGGTAACCTG TTCGCTCGCC GGTGGATTC TCGGAAACC CGAGTGTGA TTGCTTACT CAAGGAATC  
 CGAAGAATT TCCCTAGCGT GGGGACGTTG ACCCATGGAC ARGCCACGG CCGACCTAAG AGCCTTTGGG GCTCACACTT GACGAATAGA GTTTCCTTAG  
 New Cal HA

501 · N S Y I V E T P N P E N G T C Y P G Y F A D Y E E L R E Q L S S V ·  
 GTGGTCCAT ATCGTGGAGA CACCGAACC GSAGAATGG ACGTCTACC CTGTTTATT CCGAGACTAC GAAGAATTC GAGAACAAT GTCTCCGTC  
 CACCAGGATA TAGCACCTCT GTGGCTTGG CCTCTTACC TGCACGATGG GACCAATAAA GCGTCTGATG CTCTTGAAG CTCTTGTGA CAGGAGGACG

New Cal HA

601 · S S F E R F E I F P K E S S W P N H T V T G V S A S C S H N G K S S ·  
 AGTCGTTTC AGCGATTGCA AATCTTCCG AAAGAGTCAT CGTGGCCGAA TCACACGATA ACGGTTGTGT CCGCCTCATG TAGCCATAAT GGAAGTCTT  
 TCGAGCAGCA TCGTAAGCT TTAGAAAGGC TTTCTCAGTA GCACCCGCTT AGTGTGCCAT TGCCACACA GCGCAGTAC ATCGTATTA CCTTTCAGGA  
 New Cal HA

701 · F Y R N L L W L T G K N G I Y P N L S K S Y V N N K E K E V L V L ·  
 CGTCTATCG CAACCTGTG TGGCTTACTG GGA AAAACGG GTTGTACCT AATCTCAGCA AGAGCTACGT CAATAACAAA GAAAAGAGG TGCTGGTCTT  
 GCAAGATAGC GTTGGACAC ACCGAATGAC CCTTTTGGC CAACATGGA TTAGAGTCT TCTGATGCA GTTATTGTT CTTTTCTCC ACGACCAGAA  
 New Cal HA

801 · W G V H H P P N I G N Q R A L Y H T E N A Y V S V V S S H Y S R R ·  
 GTGGGGTGT CATCACCCAC CTAACATGG GAATCAGAGG GCACTGTACC AACTGAGAA TGCATACGTG AGCGTGGTGT CGAGCCACTA TAGCCGAGA  
 CACCCACAC GIATGGTGG GATTGTAACC CTTAGTCTCC CGTGACATGG TGTACTCTT ACGTATGCAC TCGCACACA GCTCGGTGAT ATCGGCTCT  
 New Cal HA

901 · F T P E I A K R P K V R D Q E G R I N Y Y W T L L E P G D T I I F E ·  
 TTCACACCAG AGATTGCGAA GCGGCCAAA GTCCCGACC AGGAGGGGGG GAFFAATAC TACTGACCC TCCTCGACC TGGCATACG ATCATCTTTG  
 AAGTGTGTC TCTAACGCTT CCGCGGTTT CAGGCGTGG TCCTCCCGG CTAATTGATG ATGACCTGG AGGAGCTCG ACCGCTATG TAGTAGAAR  
 New Cal HA

1001 · A N G N L I A P W Y A F A L S R G F G S G I I T S N A P M D E C D ·  
 AAGGAATGG TATCTTATC GCCCGTGGT ATGCTTTGC GCTTTCAGA GGATTTGGAT CAGGATCAT CACATCAAT GCGCCGATG ACGAGTGG

TTGCGTTACC ATTAGAATAG CGGGGCACCA TACGAAAACG CGAAAGTTCT CCTAAACCTA GTCCCTAGTA GTGTAGTTTA CGCGGCTACC TGCTCACGCT  
 New Cal HA  
 1101 · A R C Q T P Q G A I N S S L P F Q N V H P V T I G E C P K Y V R S  
 TGCTAAGTGT CAGACTCCCC AAGGCGCTAT CAACTCGTCG TTGCCCTTTC AAAACGTGCA CCCCCTAACG ATCGGAGAGT GTCCCAAGTA TGTCAAGATCG  
 ACGATTACCA GTCTGAGGGG TTCCGCGATA GTTGAGGAGC AACGGGAAAG TTTTCCACGT GGGGCATTGC TAGCCTCTCA CAGGGTTTAT ACAGTCTAGC  
 New Cal HA  
 1201 · A K L R M V T G L R N I P S I Q S R G L F G A I A G F I E G G W T G  
 GCGAAACITTA GGATGGTGAC CGGACTCCGC AATATCCCCT CGATCCAGTC ACGGGGATTG TTTGGAGCCA TTGGGGGCTT CATCGAAGGG GGCTGGACTG  
 CGTTTCAAT CCTACCACCTG GCCTGAGGCG TTATAGGGGA GCTAGGTAGT TGCCCTTAC AACCTCGGT AACCCCCGA GTAGCTTCCC CCGACTGAC  
 New Cal HA  
 1301 · M V D G W Y G Y H H O N E Q G S G Y A A D Q K S T Q N A I N G I T  
 CAATGGTCTGA TGGTGGTATC GGTATTCACC ACCGAAATGA GCAGGGTTCC GGGTATGCGG CGGATCAGAA ATCGACACAG AACGCAATCA ACGGGATTAC  
 CTTACCAGCT ACCCACCATG CCAATAGTGG TGGTCTTACT GTCCCAAGG CCGATACGGC GCCTAGTCTT TAGCTGTGT TTGCGTTAGT TGCCCTAATG  
 New Cal HA  
 1401 · N K V N S V I E K M N T Q F T A V G K E F N K L E R R M E N L N K  
 GAACAAGGTA AACAGCGTCA TTGAGAAGAT GAATACACAG TTTACAGCCG TGGGGAAAGA ATTCAACAAA CTCGAGCGCC GGATGGAGAA TTTGAATAG  
 CTTGTCCAT TTGTCCAGT AACTCTTCTA CTTATGTGTC AAATGTCCGC ACCCCTTCT TAAGTTGTTT GAGCTCGCGG CCTACCTCTT AAACCTATTC  
 New Cal HA  
 1501 · K V D D G F L D I W T Y N A E L L V L L E N E R T L D F H D S N V K  
 AAAGTGGACG ATGGTTTCCCT CGATATCTGG ACGTACAATG CCGAGCTGCT TGCTCTGCTC GAAAATGAGA GGACGCTCGA CTTTCATGAC TCCAATGTA  
 TTPCACCTGC TACCAAAAGGA GCTATACACC TGCAATGTTAC GCCTCGAGCA ACAGGACGAG CTTTACTCTT CCTGCGAGCT GAAAGTACTG AGTTTACT  
 New Cal HA  
 1601 · N L Y E K V K S Q L K N N A K E I G N G C F E F Y H K C N N E C M  
 AGAACCTTTA CGAGAAGTG AASTCCCAAT TGAAGAAATA CGCCAAGGAA ATTTGGAACG GCTGCTTCCA ATTCTACCCAC AAATGCAACA ATGAGTGCAT  
 TCTTGAAAT GCTCTCCAC TTCAGGGTTA ACTTCTTATT GCGGTTCCTT TAACCTTIGC CGACGAAGCT TAAGATGGTG TTTACGTTGT TACTCACGTA  
 New Cal HA  
 1701 · E S V K N G T Y D Y P K Y S E E S K L N R E K I D G V K L E S M G  
 GGAATCGGTC AAAAATGGAA CATATGATTA TCCCAATAC TCGGAGGAGT CAAGCTTAA TAGGGACAAA ATTGATGGGG TAAAACITGA GAGCATGGGT  
 CCTTAGCCAG TTTTACCTT GTRTACTAAT AGGGTTTATG AGCCTCCCA GTTTCGAATT ATCCCTCTTT TACTACCCG ATTTTGAAGT CTCGTACCCA  
 New Cal HA  
 1801 · V Y Q I L A I Y S T V A S S L V L L V S L G A I S F W M C S N G S L  
 GTATATCAGA TCTTGGCAAT CTACTCAACC GTGGCGTCTG CACTGGTACT CCTCGTCTCC CIGGGCCCA TTAGCTTTTG GATGTGTTCG AATGGATCGC  
 CATATAGTCT AGGACCGITA GATGAGTTGG CACCGCAGCA GTGACCATGA GGACACACAG GACCCGCGGT AATGAAAAC CTACACAAGC TTACCTAGCC  
 FMDV2A TM of WN E split  
 New Cal HA prE/NS1 Signal  
 1901 · Q C R I C E N F D L L K L A G D V E S N P G P A R D R S I A L T F  
 TCCAGTCCCG CATCTGCATC AACTTGCACC TGCTCAAGCT CSCGGGTGAC GTCGATCCA ACCCAGGGCC AGCCCGGGAC AGAAGCATTG CGCTCACTTT  
 AGGTCACGGC GTAGACGTAG TTGAACTGG ACGACTTCA GCGCCCACTG CAGCTTAGST TGGCTCCCG NS1  
 TM of WN E split

```

 · L A V G G V L L F L S V N V H A D T G C A I D I S R Q E L R C G S
2001 TCTCGCGGTA GAGGTTGTGC TGTGTTCCT CTCAGTGAAC GTCCACGCAG ACACGGGATG CCGGATTGAT ATCTCCAGAC AAGAATTGAG GTGCGGGTGG
 AGAGCCCAT CCTCCACAG ACAACAAGGA CAGTCACTTG CAGGTGGGTC TGTGCCCTAC GCGTAACTA TAGAGGCTG TTCTTAATC CACGCCCAGC
 NS1

 G V F I H N D V E A W M D R Y K Y Y P E T P Q G L A K I I Q K A H K ·
2101 GGGTCTTTA TCCATAACGA CGTGGAGGCG TGGATGGACA GGTATAAGTA TTACCCTGAA ACGCCGCAGG GACTTGCAGAA AATCATTGAG AAAGCCATA
 CCCCAGAAAT AGGTATTGCT GCACCTCCGC ACCTACCTCT CCATATTGAT AATGGGACTT TCGGGCGTCC CTGAACGCIT TTAGTAAGTC TTTCGGGTAT
 NS1

 · E G V C G L R S V S R L E H Q M W E A V K D E L N T L L K
2201 AGGAAGGTGT GTGTGGATG AGATCAGTCT CACGCCTTGA GCACCAGATG TGGGAGGCTG TCAAGGATGA ATTGAACACA CTTTGAAG
 TCCTCCACA CACACCTAAC TCTAGTCAGA GTCCGGAAT CGTGGTCTAC ACCCTCCGAC AGTTCCTACT TAACTTGTGT GAAACTTC

```

## OTHER EMBODIMENTS

All publications, patent applications, and patents mentioned in this specification are incorporated herein by reference in their entirety as if each individual publication, patent application, or patent were specifically and individually indicated to be incorporated by reference. In particular, U.S. 2011/0135686 is hereby incorporated by reference in its entirety.

Various modifications and variations of the described viruses, vectors, compositions, and methods of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the inven-

tion has been described in connection with specific embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the fields of medicine, pharmacology, or related fields are intended to be within the scope of the invention. Use of singular forms herein, such as "a" and "the," does not exclude indication of the corresponding plural form, unless the context indicates to the contrary. Similarly, use of plural terms does not exclude indication of a corresponding singular form. Other embodiments are within the scope of the following claims.

## SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 78

<210> SEQ ID NO 1  
<211> LENGTH: 12  
<212> TYPE: PRT  
<213> ORGANISM: Yellow Fever Virus

<400> SEQUENCE: 1

Ser His Asp Val Leu Thr Val Gln Phe Leu Ile Leu  
1 5 10

<210> SEQ ID NO 2  
<211> LENGTH: 8  
<212> TYPE: PRT  
<213> ORGANISM: Tick-borne Encephalitis Virus

<400> SEQUENCE: 2

Gly Met Leu Gly Met Thr Ile Ala  
1 5

<210> SEQ ID NO 3  
<211> LENGTH: 20  
<212> TYPE: PRT  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 3

Ser His Asp Val Leu Thr Val Gln Phe Leu Ile Leu Gly Met Leu Gly  
1 5 10 15

Met Thr Ile Ala  
20

<210> SEQ ID NO 4  
<211> LENGTH: 20  
<212> TYPE: PRT  
<213> ORGANISM: Tick-borne Encephalitis Virus

<400> SEQUENCE: 4

Gly Gly Thr Asp Trp Met Ser Trp Leu Leu Val Ile Gly Met Leu Gly  
1 5 10 15

Met Thr Ile Ala  
20

<210> SEQ ID NO 5  
<211> LENGTH: 9  
<212> TYPE: PRT  
<213> ORGANISM: Borrelia burgdorferi

-continued

&lt;400&gt; SEQUENCE: 5

Tyr Val Leu Glu Gly Thr Leu Thr Ala  
 1 5

&lt;210&gt; SEQ ID NO 6

&lt;211&gt; LENGTH: 9

&lt;212&gt; TYPE: PRT

<213> ORGANISM: *Borrelia afzelii*

&lt;400&gt; SEQUENCE: 6

Phe Thr Leu Glu Gly Lys Val Ala Asn  
 1 5

&lt;210&gt; SEQ ID NO 7

&lt;211&gt; LENGTH: 9

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 7

Phe Thr Leu Glu Gly Lys Leu Thr Ala  
 1 5

&lt;210&gt; SEQ ID NO 8

&lt;211&gt; LENGTH: 273

&lt;212&gt; TYPE: PRT

<213> ORGANISM: *Borrelia burgdorferi*

&lt;400&gt; SEQUENCE: 8

Met Lys Lys Tyr Leu Leu Gly Ile Gly Leu Ile Leu Ala Leu Ile Ala  
 1 5 10 15

Cys Lys Gln Asn Val Ser Ser Leu Asp Glu Lys Asn Ser Val Ser Val  
 20 25 30

Asp Leu Pro Gly Glu Met Lys Val Leu Val Ser Lys Glu Lys Asn Lys  
 35 40 45

Asp Gly Lys Tyr Asp Leu Ile Ala Thr Val Asp Lys Leu Glu Leu Lys  
 50 55 60

Gly Thr Ser Asp Lys Asn Asn Gly Ser Gly Val Leu Glu Gly Val Lys  
 65 70 75 80

Ala Asp Lys Ser Lys Val Lys Leu Thr Ile Ser Asp Asp Leu Gly Gln  
 85 90 95

Thr Thr Leu Glu Val Phe Lys Glu Asp Gly Lys Thr Leu Val Ser Lys  
 100 105 110

Lys Val Thr Ser Lys Asp Lys Ser Ser Thr Glu Glu Lys Phe Asn Glu  
 115 120 125

Lys Gly Glu Val Ser Glu Lys Ile Ile Thr Arg Ala Asp Gly Thr Arg  
 130 135 140

Leu Glu Tyr Thr Gly Ile Lys Ser Asp Gly Ser Gly Lys Ala Lys Glu  
 145 150 155 160

Val Leu Lys Gly Tyr Val Leu Glu Gly Thr Leu Thr Ala Glu Lys Thr  
 165 170 175

Thr Leu Val Val Lys Glu Gly Thr Val Thr Leu Ser Lys Asn Ile Ser  
 180 185 190

Lys Ser Gly Glu Val Ser Val Glu Leu Asn Asp Thr Asp Ser Ser Ala  
 195 200 205

Ala Thr Lys Lys Thr Ala Ala Trp Asn Ser Gly Thr Ser Thr Leu Thr  
 210 215 220

-continued

---

```
Ile Thr Val Asn Ser Lys Lys Thr Lys Asp Leu Val Phe Thr Lys Glu
225 230 235 240
```

```
Asn Thr Ile Thr Val Gln Gln Tyr Asp Ser Asn Gly Thr Lys Leu Glu
 245 250 255
```

```
Gly Ser Ala Val Glu Ile Thr Lys Leu Asp Glu Ile Lys Asn Ala Leu
 260 265 270
```

Lys

```
<210> SEQ ID NO 9
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: X is Glu or Gly
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (5)..(5)
<223> OTHER INFORMATION: X is Met or Ile
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (6)..(6)
<223> OTHER INFORMATION: X is Lys, Thr or Gly
```

&lt;400&gt; SEQUENCE: 9

```
Leu Pro Gly Xaa Xaa Xaa Val Leu
1 5
```

```
<210> SEQ ID NO 10
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (6)..(6)
<223> OTHER INFORMATION: X is Asn, Ser, or Asp
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (11)..(11)
<223> OTHER INFORMATION: X is Val, or Thr
```

&lt;400&gt; SEQUENCE: 10

```
Gly Thr Ser Asp Lys Xaa Asn Gly Ser Gly Xaa
1 5 10
```

```
<210> SEQ ID NO 11
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (1)..(1)
<223> OTHER INFORMATION: Xaa is Asn, His, or Glu
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (3)..(3)
<223> OTHER INFORMATION: Xaa is Ser, Pro, Leu, Ala, or Ser
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: Xaa is Lys or Asn
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (8)..(8)
```

-continued

---

```

<223> OTHER INFORMATION: Xaa is Val or Ile
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (9)..(9)
<223> OTHER INFORMATION: Xaa is Ser or Thr
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (10)..(10)
<223> OTHER INFORMATION: Xaa is Val or Ala
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (11)..(11)
<223> OTHER INFORMATION: Xaa is Glu or Ala
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (13)..(13)
<223> OTHER INFORMATION: Xaa is Asn or Asp
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (15)..(15)
<223> OTHER INFORMATION: Xaa is Thr or Ser
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (16)..(16)
<223> OTHER INFORMATION: Xaa is Asp or Asn
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (17)..(17)
<223> OTHER INFORMATION: Xaa is Ser or Thr
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (18)..(18)
<223> OTHER INFORMATION: Xaa is Ser or Thr
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (19)..(19)
<223> OTHER INFORMATION: Xaa is Ala, Gln, or Arg
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (25)..(25)
<223> OTHER INFORMATION: Xaa is Ala or Gly
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (26)..(26)
<223> OTHER INFORMATION: Xaa is Ala, Lys, or Thr
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (28)..(28)
<223> OTHER INFORMATION: Xaa is Asn or Asp
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (29)..(29)
<223> OTHER INFORMATION: Xaa is Ser or Ala
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (30)..(30)
<223> OTHER INFORMATION: Xaa is Gly, Asn, or Lys

```

```

<400> SEQUENCE: 11

```

```

Xaa Ile Xaa Xaa Ser Gly Glu Xaa Xaa Xaa Xaa Leu Xaa Asp Xaa Xaa
1 5 10 15

```

```

Xaa Xaa Xaa Ala Thr Lys Lys Thr Xaa Xaa Trp Xaa Xaa Xaa Thr
20 25 30

```

```

<210> SEQ ID NO 12
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (2)..(2)
<223> OTHER INFORMATION: Xaa is Asn or Ala
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE

```

-continued

---

```

<222> LOCATION: (5)..(5)
<223> OTHER INFORMATION: Xaa is Lys or Asn
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (9)..(9)
<223> OTHER INFORMATION: Xaa is Ser, Asn, Lys, or Thr
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (14)..(14)
<223> OTHER INFORMATION: Xaa is Thr or Lys
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (15)..(15)
<223> OTHER INFORMATION: Xaa is Thr or Lys
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (17)..(17)
<223> OTHER INFORMATION: Xaa is Asp or Lys
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (19)..(19)
<223> OTHER INFORMATION: Xaa is Ile or Leu

```

```

<400> SEQUENCE: 12

```

```

Ser Xaa Gly Thr Xaa Leu Glu Gly Xaa Ala Val Glu Ile Xaa Xaa Leu
1 5 10 15

```

```

Xaa Glu Xaa Lys Asn
20

```

```

<210> SEQ ID NO 13
<211> LENGTH: 154
<212> TYPE: PRT
<213> ORGANISM: Rhipicephalus appendiculatus

```

```

<400> SEQUENCE: 13

```

```

Met Lys Ala Phe Phe Val Leu Ser Leu Leu Ser Thr Ala Ala Leu Thr
1 5 10 15

```

```

Asn Ala Ala Arg Ala Gly Arg Leu Gly Ser Asp Leu Asp Thr Phe Gly
20 25 30

```

```

Arg Val His Gly Asn Leu Tyr Ala Gly Ile Glu Arg Ala Gly Pro Arg
35 40 45

```

```

Gly Tyr Pro Gly Leu Thr Ala Ser Ile Gly Gly Glu Val Gly Ala Arg
50 55 60

```

```

Leu Gly Gly Arg Ala Gly Val Gly Val Ser Ser Tyr Gly Tyr Gly Tyr
65 70 75 80

```

```

Pro Ser Trp Gly Tyr Pro Tyr Gly Gly Tyr Gly Gly Tyr Gly Tyr
85 90 95

```

```

Gly Gly Tyr Gly Gly Tyr Asp Gln Gly Phe Gly Ser Ala Tyr Gly Gly
100 105 110

```

```

Tyr Pro Gly Tyr Tyr Gly Tyr Tyr Tyr Pro Ser Gly Tyr Gly Gly Gly
115 120 125

```

```

Tyr Gly Gly Ser Tyr Gly Gly Ser Tyr Gly Gly Ser Tyr Thr Tyr Pro
130 135 140

```

```

Asn Val Arg Ala Ser Ala Gly Ala Ala Ala
145 150

```

```

<210> SEQ ID NO 14
<211> LENGTH: 184
<212> TYPE: PRT
<213> ORGANISM: Ixodes scapularis

```

```

<400> SEQUENCE: 14

```

```

Met Arg Thr Ala Phe Thr Cys Ala Leu Leu Ala Ile Ser Phe Leu Gly
1 5 10 15

```

-continued

---

Ser Pro Cys Ser Ser Ser Glu Asp Gly Leu Glu Gln Asp Thr Ile Val  
                   20                                  25                                  30  
 Glu Thr Thr Thr Gln Asn Leu Tyr Glu Arg His Tyr Arg Asn His Ser  
                   35                                  40                                  45  
 Gly Leu Cys Gly Ala Gln Tyr Arg Asn Ser Ser His Ala Glu Ala Val  
                   50                                  55                                  60  
 Tyr Asn Cys Thr Leu Asn His Leu Pro Pro Val Val Asn Ala Thr Trp  
                   65                                  70                                  75                                  80  
 Glu Gly Ile Arg His Arg Ile Asn Lys Thr Ile Pro Gln Phe Val Lys  
                   85                                  90                                  95  
 Leu Ile Cys Asn Phe Thr Val Ala Met Pro Gln Glu Phe Tyr Leu Val  
                   100                                  105                                  110  
 Tyr Met Gly Ser Asp Gly Asn Ser Asp Phe Glu Glu Asp Lys Glu Ser  
                   115                                  120                                  125  
 Thr Gly Thr Asp Glu Asp Ser Asn Thr Gly Ser Ser Ala Ala Ala Lys  
                   130                                  135                                  140  
 Val Thr Glu Ala Leu Ile Ile Glu Ala Glu Glu Asn Cys Thr Ala His  
                   145                                  150                                  155                                  160  
 Ile Thr Gly Trp Thr Thr Glu Thr Pro Thr Thr Leu Glu Pro Thr Thr  
                   165                                  170                                  175  
 Glu Ser Gln Phe Glu Ala Ile Pro  
                   180

<210> SEQ ID NO 15  
 <211> LENGTH: 177  
 <212> TYPE: PRT  
 <213> ORGANISM: Ixodes scapularis

<400> SEQUENCE: 15

Met Arg Thr Ala Leu Thr Cys Ala Leu Leu Ala Ile Ser Phe Leu Gly  
 1                  5                                  10                                  15  
 Ser Pro Cys Ser Ser Ser Glu Gly Gly Leu Glu Lys Asp Ser Arg Val  
                   20                                  25                                  30  
 Glu Thr Thr Thr Gln Asn Leu Tyr Glu Arg Tyr Tyr Arg Lys His Pro  
                   35                                  40                                  45  
 Gly Leu Cys Gly Ala Gln Tyr Arg Asn Ser Ser His Ala Glu Ala Val  
                   50                                  55                                  60  
 Tyr Asn Cys Thr Leu Ser Leu Leu Pro Leu Ser Val Asn Thr Thr Trp  
                   65                                  70                                  75                                  80  
 Glu Gly Ile Arg His Arg Ile Asn Lys Thr Ile Pro Glu Phe Val Asn  
                   85                                  90                                  95  
 Leu Ile Cys Asn Phe Thr Val Ala Met Pro Asp Gln Phe Tyr Leu Val  
                   100                                  105                                  110  
 Tyr Met Gly Ser Asn Gly Asn Ser Tyr Ser Glu Glu Asp Glu Asp Gly  
                   115                                  120                                  125  
 Lys Thr Gly Ser Ser Ala Ala Val Gln Val Thr Glu Gln Leu Ile Ile  
                   130                                  135                                  140  
 Gln Ala Glu Glu Asn Cys Thr Ala His Ile Thr Gly Trp Thr Thr Glu  
                   145                                  150                                  155                                  160  
 Ala Pro Thr Thr Leu Glu Pro Thr Thr Glu Thr Gln Phe Glu Ala Ile  
                   165                                  170                                  175

Ser

<210> SEQ ID NO 16

-continued

---

<211> LENGTH: 19  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza A virus

<400> SEQUENCE: 16

Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly  
 1 5 10 15

Phe Leu Glu

<210> SEQ ID NO 17  
 <211> LENGTH: 23  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza A virus

<400> SEQUENCE: 17

Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly  
 1 5 10 15

Phe Leu Glu Gly Ser Gly Cys  
 20

<210> SEQ ID NO 18  
 <211> LENGTH: 19  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza B virus

<400> SEQUENCE: 18

Asn Asn Ala Thr Phe Asn Tyr Thr Asn Val Asn Pro Ile Ser His Ile  
 1 5 10 15

Arg Gly Ser

<210> SEQ ID NO 19  
 <211> LENGTH: 24  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza A virus

<400> SEQUENCE: 19

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly  
 1 5 10 15

Cys Arg Cys Asn Asp Ser Ser Asp  
 20

<210> SEQ ID NO 20  
 <211> LENGTH: 24  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza A virus

<400> SEQUENCE: 20

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu  
 1 5 10 15

Cys Arg Cys Ser Asp Ser Ser Asp  
 20

<210> SEQ ID NO 21  
 <211> LENGTH: 24  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza A virus

<400> SEQUENCE: 21

Met Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Gly  
 1 5 10 15

Cys Arg Cys Ser Asp Ser Ser Asp  
 20

-continued

<210> SEQ ID NO 22  
 <211> LENGTH: 8  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza A virus

<400> SEQUENCE: 22

Glu Val Glu Thr Pro Thr Arg Asn  
 1 5

<210> SEQ ID NO 23  
 <211> LENGTH: 23  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza A virus

<400> SEQUENCE: 23

Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Cys  
 1 5 10 15  
 Arg Cys Asn Asp Ser Ser Asp  
 20

<210> SEQ ID NO 24  
 <211> LENGTH: 17  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza A virus

<400> SEQUENCE: 24

Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Cys  
 1 5 10 15  
 Arg

<210> SEQ ID NO 25  
 <211> LENGTH: 6  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza B virus

<400> SEQUENCE: 25

Met Leu Glu Pro Phe Gln  
 1 5

<210> SEQ ID NO 26  
 <211> LENGTH: 12  
 <212> TYPE: PRT  
 <213> ORGANISM: Influenza B virus

<400> SEQUENCE: 26

Leu Glu Pro Phe Gln Ile Leu Ser Ile Ser Gly Cys  
 1 5 10

<210> SEQ ID NO 27  
 <211> LENGTH: 24  
 <212> TYPE: PRT  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Avian Influenza A virus Subtype H5N1

<400> SEQUENCE: 27

Met Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Gly  
 1 5 10 15  
 Cys Arg Cys Ser Asp Ser Ser Asp  
 20

<210> SEQ ID NO 28

-continued

---

```

<211> LENGTH: 28
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 28

Arg Lys Arg Arg Ser His Asp Val Leu Thr Val Gln Phe Leu Ile Leu
 1 5 10 15

Gly Met Leu Gly Met Thr Ile Ala Ala Thr Val Arg
 20 25

<210> SEQ ID NO 29
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 29

aggaaacgcc gttcccatga tgttctgact gtgcaattcc taattttggg catgctgggc 60
atgacaatcg cagctacggg tcgc 84

<210> SEQ ID NO 30
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 30

tcctttgcgg caagggtact acaagactga cacgttaagg attaaaaccc gtacgacctg 60
tactgttagc gtcgatgcca agcg 84

<210> SEQ ID NO 31
<211> LENGTH: 28
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 31

Arg Lys Arg Arg Ser His Asp Val Leu Thr Val Gln Phe Leu Ile Leu
 1 5 10 15

Gly Met Leu Ala Cys Val Gly Ala Ala Thr Val Arg
 20 25

<210> SEQ ID NO 32
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 32

aggaaacgcc gttcccatga tgttctgact gtgcaattcc taattttggg catgctggct 60
tgtgtcggag cagctaccgt gcca 84

<210> SEQ ID NO 33
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

-continued

&lt;400&gt; SEQUENCE: 33

```
tcctttgctg caagggtact acaagactga cactgtaagg attaaaaacc gtacgaccga 60
acacagcctc gtcgatggca cgct 84
```

&lt;210&gt; SEQ ID NO 34

&lt;211&gt; LENGTH: 28

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 34

```
Gln Lys Lys Arg Gly Gly Thr Asp Trp Met Ser Trp Leu Leu Val Ile
 1 5 10 15
Gly Met Leu Gly Met Thr Ile Ala Ala Thr Val Arg
 20 25
```

&lt;210&gt; SEQ ID NO 35

&lt;211&gt; LENGTH: 84

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 35

```
caaaagaaac gggggggaac agactggatg agctggctgc tcgtaatcgg catgctgggc 60
atgacaatcg cagctacggg tcgc 84
```

&lt;210&gt; SEQ ID NO 36

&lt;211&gt; LENGTH: 84

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 36

```
gtttttcttg ccccccttg tctgacctac tcgaccgacg agcattagcc gtacgaccgg 60
tactgttagc gtcgatgccg agcg 84
```

&lt;210&gt; SEQ ID NO 37

&lt;211&gt; LENGTH: 26

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 37

```
Gln Lys Lys Arg Gly Gly Lys Thr Gly Ile Ala Val Met Ile Gly Met
 1 5 10 15
Leu Ala Cys Val Gly Ala Ala Thr Val Arg
 20 25
```

&lt;210&gt; SEQ ID NO 38

&lt;211&gt; LENGTH: 78

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 38

```
caaaagaaac gcgggggaaa gacaggcata gctgtgatga taggcatgct ggcttgtgtc 60
```

-continued

ggagcagcta ccgtgcga

78

<210> SEQ ID NO 39  
 <211> LENGTH: 78  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 39

gttttctttg cgcccccttt ctgtccgat cgacactact atccgtacga ccgaacacag 60

cctcgtcgat ggacgct 78

<210> SEQ ID NO 40  
 <211> LENGTH: 793  
 <212> TYPE: PRT  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 40

Met Ser Gly Arg Lys Ala Gln Gly Lys Thr Leu Gly Val Asn Met Val  
1 5 10 15Arg Arg Gly Val Arg Ser Leu Ser Asn Lys Ile Lys Gln Lys Thr Lys  
20 25 30Gln Ile Gly Asn Arg Pro Gly Pro Ser Arg Gly Val Gln Gly Phe Ile  
35 40 45Phe Phe Phe Leu Phe Asn Ile Leu Thr Gly Lys Lys Ile Thr Ala His  
50 55 60Leu Lys Arg Leu Trp Lys Met Leu Asp Pro Arg Gln Gly Leu Ala Val  
65 70 75 80Leu Arg Lys Val Lys Arg Val Val Ser Leu Met Arg Gly Leu Ser Ser  
85 90 95Arg Lys Arg Arg Ser His Asp Val Leu Thr Val Gln Phe Leu Ile Leu  
100 105 110Gly Met Leu Gly Met Thr Ile Ala Ala Thr Val Arg Lys Glu Arg Asp  
115 120 125Gly Ser Thr Val Ile Arg Ala Glu Gly Lys Asp Ala Ala Thr Gln Val  
130 135 140Arg Val Glu Asn Gly Thr Cys Val Ile Leu Ala Thr Asp Met Gly Ser  
145 150 155 160Trp Cys Asp Asp Ser Leu Ser Tyr Glu Cys Val Thr Ile Asp Gln Gly  
165 170 175Glu Glu Pro Val Asp Val Asp Cys Phe Cys Arg Asn Val Asp Gly Val  
180 185 190Tyr Leu Glu Tyr Gly Arg Cys Gly Lys Gln Glu Gly Ser Arg Thr Arg  
195 200 205Arg Ser Val Leu Ile Pro Ser His Ala Gln Gly Glu Leu Thr Gly Arg  
210 215 220Gly His Lys Trp Leu Glu Gly Asp Ser Leu Arg Thr His Leu Thr Arg  
225 230 235 240Val Glu Gly Trp Val Trp Lys Asn Arg Leu Leu Ala Leu Ala Met Val  
245 250 255Thr Val Val Trp Leu Thr Leu Glu Ser Val Val Thr Arg Val Ala Val  
260 265 270Leu Val Val Leu Leu Cys Leu Ala Pro Val Tyr Ala Ser Arg Cys Thr  
275 280 285

-continued

---

His Leu Glu Asn Arg Asp Phe Val Thr Gly Thr Gln Gly Thr Thr Arg  
 290 295 300  
 Val Thr Leu Val Leu Glu Leu Gly Gly Cys Val Thr Ile Thr Ala Glu  
 305 310 315 320  
 Gly Lys Pro Ser Met Asp Val Trp Leu Asp Ala Ile Tyr Gln Glu Asn  
 325 330 335  
 Pro Ala Gln Thr Arg Glu Tyr Cys Leu His Ala Lys Leu Ser Asp Thr  
 340 345 350  
 Lys Val Ala Ala Arg Cys Pro Thr Met Gly Pro Ala Thr Leu Ala Glu  
 355 360 365  
 Glu His Gln Gly Gly Thr Val Cys Lys Arg Asp Gln Ser Asp Arg Gly  
 370 375 380  
 Trp Gly Asn His Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala Cys  
 385 390 395 400  
 Val Lys Ala Ala Cys Glu Ala Lys Lys Lys Ala Thr Gly His Val Tyr  
 405 410 415  
 Asp Ala Asn Lys Ile Val Tyr Thr Val Lys Val Glu Pro His Thr Gly  
 420 425 430  
 Asp Tyr Val Ala Ala Asn Glu Thr His Ser Gly Arg Lys Thr Ala Ser  
 435 440 445  
 Phe Thr Val Ser Ser Glu Lys Thr Ile Leu Thr Met Gly Glu Tyr Gly  
 450 455 460  
 Asp Val Ser Leu Leu Cys Arg Val Ala Ser Gly Val Asp Leu Ala Gln  
 465 470 475 480  
 Thr Val Ile Leu Glu Leu Asp Lys Thr Val Glu His Leu Pro Thr Ala  
 485 490 495  
 Trp Gln Val His Arg Asp Trp Phe Asn Asp Leu Ala Leu Pro Trp Lys  
 500 505 510  
 His Glu Gly Ala Arg Asn Trp Asn Asn Ala Glu Arg Leu Val Glu Phe  
 515 520 525  
 Gly Ala Pro His Ala Val Lys Met Asp Val Tyr Asn Leu Gly Asp Gln  
 530 535 540  
 Thr Gly Val Leu Leu Lys Ala Leu Ala Gly Val Pro Val Ala His Ile  
 545 550 555 560  
 Glu Gly Thr Lys Tyr His Leu Lys Ser Gly His Val Thr Cys Glu Val  
 565 570 575  
 Gly Leu Glu Lys Leu Lys Met Lys Gly Leu Thr Tyr Thr Met Cys Asp  
 580 585 590  
 Lys Thr Lys Phe Thr Trp Lys Arg Ala Pro Thr Asp Ser Gly His Asp  
 595 600 605  
 Thr Val Val Met Glu Val Thr Phe Ser Gly Thr Lys Pro Cys Arg Ile  
 610 615 620  
 Pro Val Arg Ala Val Ala His Gly Ser Pro Asp Val Asn Val Ala Met  
 625 630 635 640  
 Leu Ile Thr Pro Asn Pro Thr Ile Glu Asn Asn Gly Gly Gly Phe Ile  
 645 650 655  
 Glu Met Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr Val Gly Glu Leu  
 660 665 670  
 Ser Tyr Gln Trp Phe Gln Lys Gly Ser Ser Ile Gly Arg Val Phe Gln  
 675 680 685  
 Lys Thr Lys Lys Gly Ile Glu Arg Leu Thr Val Ile Gly Glu His Ala  
 690 695 700

-continued

Trp Asp Phe Gly Ser Ala Gly Gly Phe Leu Ser Ser Ile Gly Lys Ala  
705 710 715 720

Leu His Thr Val Leu Gly Gly Ala Phe Asn Ser Ile Phe Gly Gly Val  
725 730 735

Gly Phe Leu Pro Lys Leu Leu Leu Gly Val Ala Leu Ala Trp Leu Gly  
740 745 750

Leu Asn Met Arg Asn Pro Thr Met Ser Met Ser Phe Leu Leu Ala Gly  
755 760 765

Val Leu Val Leu Ala Met Thr Leu Gly Val Gly Ala Asp Gln Gly Cys  
770 775 780

Ala Ile Asn Phe Gly Lys Arg Glu Leu  
785 790

<210> SEQ ID NO 41  
<211> LENGTH: 2500  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 41

```

agtaaatcct gtgtgctaata tgagggtgcat tgggtctgcaa atcgagttgc taggcaataa 60
acacatttgg attaatttta atcgttcgtt gagcgattag cagagaactg accagaacat 120
gtctggctcg aaagctcagg gaaaaaccct gggcgtaaat atggtacgac gaggagtctg 180
ctccttgta aacaaaataa aacaaaaaac aaaacaaatt ggaacagac ctggaccttc 240
aagaggtgtt caaggattta tctttttctt tttgttcaac attttgactg gaaaaaagat 300
cacagcccac ctaaagaggt tgtgaaaaat gctggaccca agacaaggct tggctgttct 360
aaggaaagtc aagagagtgg tggccagtt gatgagagga ttgtcctcaa ggaacgccg 420
ttcccatgat gttctgactg tgcaattcct aattttgggc atgctgggca tgacaatcgc 480
agctacggtt cgcaaggaaa gagacggcag tacggtcata cgcgcggaag gtaaggatgc 540
cgctacccaa gtgagagtgg aaaatggtag ctgctcatt ctggccaccg acatgggctc 600
ttggtgtgat gatagccttt cttatgagtg cgtaaccata gatcaagggt aggaacctgt 660
tgacgttgat tgcttctgcc gaaacgtgga tggggtgtat ctgcaatatg gacgggtgtg 720
taaacaagaa ggaagcagaa ccagacgctc agtgcttata ccctcccacg ctcaaggaga 780
gctgaccgga cggggacata aatggttga gggcgactca ctccgaacac atttgaccgg 840
cgctgagggc tgggtctgga aaaatcggct gttggccctc gctatggtga cagtcgtttg 900
gctcacgctg gagtctgtgg ttactcgcgt ggcagtgctg gtggtgctcc tctgtcttgc 960
ccctgtctac gcgtccaggt gtactcattt ggaaaacaga gattttgtca ccggcaacca 1020
ggggacgact cgggtaaccc tgggtcttga actgggtggt tgcgttacta ttaccgctga 1080
gggcaaaccc tctatggatg tgtggctgga tgcaatctat caggagaatc ccgcacaaac 1140
cagggaatat tgccttcacg caaagctgtc cgatacaaag gtcgcggtc ggtgcccaac 1200
aatgggaccg gccaccctgg cggaggaaca tcagggaggt acagtgtgca aacgggacca 1260
gagtgataga ggctggggtg atcactgcgg cctgttcggc aaaggaagta ttgtcgttg 1320
cgtaaggca gcctgtgagg ccaaaaagaa ggctactggg cacgtctatg acgccaacaa 1380
gatcgtttat acagtgaag tggaaccaca cacaggggat tacgtggcgg ccaacgagac 1440
tcattccggt cgcaaacgg ccagcttcac cgtgtcatcc gaaaagacca tcctcactat 1500
gggggagtat ggcgacgttt ctctgctctg ccgggtggct agcggagtgc acctggccca 1560

```

-continued

---

```

gacagtcac ctaggaactgg ataaaacagt tgagcatctg cctaccgctt ggcaggtgca 1620
cagggattgg tttaacgacc ttgccctgcc atggaaacat gaaggagcga gaaactggaa 1680
taatgcagag cgactcgtag aattcgggtgc cctcatgcc gtgaagatgg acgtctacaa 1740
tctgggtgat cagaccggcg ttctecttaa agctctcgtt gggtaccag ttgccacat 1800
cgaaggaacg aagtaccacc tgaagtcagg ccatgtaact tgcgaggtgg gcctggagaa 1860
gttgaataat aaaggtctta cgtacacaat gtgtgacaag accaagtcca catggaagag 1920
ggccccaca gatagcggcc acgatactgt ggtgatggag gtgacctttt ctggaacaaa 1980
accctgcaga ataccctgac gggctgtagc tcacggatct cccgatgtca atgttgctat 2040
gctgattaca cctaacccta ccatcgagaa taacgggtgt ggttttattg agatgcagct 2100
tccgccagcc gataacatca tctacgtggg cgaactctct taccagtgtt ttcagaaag 2160
gagttcaatt gggcgggtct tccaaaaaac gaagaagga atcgaacgat tgacggttat 2220
cggcgagcac gcattgggatt ttggttccgc agggggattc ctgtcttcta ttggtaaggc 2280
actgcatacc gtgctggggg ggcattcaa ttctatttcc gggggcgttg ggttccctgc 2340
taaactcctg ctgggagtag ccttgctctg gttgggactg aatatgcgga atccgacgat 2400
gtccatgtca ttctcttgg ccggcgtgct tgtactggcc atgacactgg gcgttggcgc 2460
cgatcaagga tgcgccatca actttggcaa gagagagctc 2500

```

```

<210> SEQ ID NO 42
<211> LENGTH: 2496
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 42

```

```

tcatttagga cacacgatta actccacgta accagacgtt tagctcaacg atccggttatt 60
tgtgtaaacc taattaaaat tagcaagcaa ctcgctaata gtctcttgac tggctctgta 120
cagaccagca tttcgagtcc ctttttggga cccgcagtta taccatgctg ctctcaagc 180
gaggaaacagt ttgttttatt ttgttttttg ttttgtttaa cctttgtctg gacctggaag 240
ttctccacaa gttcctaaat agaaaaagaa aaacaagttg taaaactgac ctttttctta 300
gtgtcgggtg gatttctoca acaccttta cgacctgggt tctgttccga accgacaaga 360
ttcttttcag ttctctcacc accggtcaaa ctactctcct aacaggagtt cctttgcggc 420
aagggtacta caagactgac acgttaagga ttaaaacccg tacgacctgt actgtagcgt 480
tcgatgcaa gcgttctctt ctctgcgctc atgccaatgt gcgcgccttc cattcctacg 540
gcgatgggtt cactctcacc ttttaccatg gacgcagtaa gaccgggtggc tgtacctgag 600
aaccacacta ctatcgaaa gaatactcac gcattggtat ctagtccac tccttggaca 660
actgcaacta acgaagacgg ctttgcacct accccacata gagcttatac ctgccacacc 720
atgtgttctt ccttctctt ggtctgcgag tcacgaatat gggaggggtgc gagttctctt 780
cgactggcct gccctgtat ttaccaacct cccgctgagt gaggcttctg taaactgggc 840
gcagctcccg acccagacct ttttagccga caaccgggag cgataccact gtcagcaaac 900
cgagtgcgac ctcagacacc aatgagcgca ccgtcacgac caccacgagg agacagaacg 960
gggacagatg cgcaggtcca catgagtaaa ccttttctct ctaaaacagt ggcggtgggt 1020
cccctgctga gccattggg accacgaact tgaccacca acgcaatgat aatggcgact 1080

```

-continued

---

```

cccgtttggg agatacctac acaccgacct acgtttagata gtcctcttag ggcgtgttt 1140
gtcccttata acggaagtgc gtttcgacag gctatgtttc cagcgccgat ccacgggtt 1200
ttaccctggc cgggtgggacc gctccttgt agtcctcca tgtcacacgt ttgccctgt 1260
ctcactatct ccgaccccat tagtgaogcc ggacaagccg tttccttcat aacagcgaac 1320
gcagttccgt cgggacactcc ggtttttctt ccgatgacct gtgcagatac tgcggttgt 1380
ctagcaaata tgtaacttcc accttggtgt gtgtcccta atgcaccgcc ggttgetctg 1440
agtaaggcca gcgttttgcc ggtcgaagtg gcacagtagg cttttctggt aggagtgata 1500
ccccctcata ccgctgcaaa gagacgagac ggcccaccga tcgcctcagc tggaccgggt 1560
ctgtcagtag gaccttgacc tattttgtca actcgtagac ggatggcgaa ccgtccactg 1620
gtccctaacc aaattgctgg aacgggacgg tacctttgta cttcctcgt ctttgacct 1680
attacgtctc gctgagcctc ttaagccacg gggagtacgg cacttctacc tgcagatgt 1740
agacccta gtctggccgc aagaggaatt tcgagagcga ccgcatggtc aacgggtgta 1800
gcttccttgc ttcattggtg acttcagtcc ggtacattga acgctccacc cggacctct 1860
caacttttac tttccagaat gcatgtgta cacactgttc tggttcaagt gtaccttctc 1920
ccgggggtgt ctatcgccg tgetatgaca ccaactctc cactggaaaa gacctgttt 1980
tgggacgtct tatgggcacg cccgacatcg agtgcctaga gggctacagt tacaacgata 2040
cgactaatgt ggattgggat ggtagctctt attgccacca ccaaaataac tctacgtcga 2100
aggcggtccg ctattgtagt agatgacccc gcttgagaga atggtcacca aagtctttcc 2160
ctcaagttaa cccgccaga aggttttttg cttcttccct tagcttgcta actgccaata 2220
gccgctctgt cgtaccctaa aaccaaggcg tcccctaag gacagaagat aaccattccg 2280
tgacgtatgg caccacccc cgcgtaagtt aagataaaag cccccgcacc ccaaggacgg 2340
atgtgaggac gacctctatc gggaccggac caacctgac ttatacgcct taggctgcta 2400
tacagtaagg agaaccggcc gcacgaacat gaccggtact gtgacccgca accgcggtc 2460
gttctacgc ggtagttgaa accgttctct ctcgag 2496

```

&lt;210&gt; SEQ ID NO 43

&lt;211&gt; LENGTH: 793

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 43

```

Met Ser Gly Arg Lys Ala Gln Gly Lys Thr Leu Gly Val Asn Met Val
1 5 10 15
Arg Arg Gly Val Arg Ser Leu Ser Asn Lys Ile Lys Gln Lys Thr Lys
20 25 30
Gln Ile Gly Asn Arg Pro Gly Pro Ser Arg Gly Val Gln Gly Phe Ile
35 40 45
Phe Phe Phe Leu Phe Asn Ile Leu Thr Gly Lys Lys Ile Thr Ala His
50 55 60
Leu Lys Arg Leu Trp Lys Met Leu Asp Pro Arg Gln Gly Leu Ala Val
65 70 75 80
Leu Arg Lys Val Lys Arg Val Val Ala Ser Leu Met Arg Gly Leu Ser
85 90 95
Ser Arg Lys Arg Arg Ser His Asp Val Leu Thr Val Gln Phe Leu Ile
100 105 110

```

-continued

---

Leu Gly Met Leu Ala Cys Val Gly Ala Ala Thr Val Arg Lys Glu Arg  
 115 120 125  
 Asp Gly Ser Thr Val Ile Arg Ala Glu Gly Lys Asp Ala Ala Thr Gln  
 130 135 140  
 Val Arg Val Glu Asn Gly Thr Cys Val Ile Leu Ala Thr Asp Met Gly  
 145 150 155 160  
 Ser Trp Cys Asp Asp Ser Leu Ser Tyr Glu Cys Val Thr Ile Asp Gln  
 165 170 175  
 Gly Glu Glu Pro Val Asp Val Asp Cys Phe Cys Arg Asn Val Asp Gly  
 180 185 190  
 Val Tyr Leu Glu Tyr Gly Arg Cys Gly Lys Gln Glu Gly Ser Arg Thr  
 195 200 205  
 Arg Arg Ser Val Leu Ile Pro Ser His Ala Gln Gly Glu Leu Thr Gly  
 210 215 220  
 Arg Gly His Lys Trp Leu Glu Gly Asp Ser Leu Arg Thr His Leu Thr  
 225 230 235 240  
 Arg Val Glu Gly Trp Val Trp Lys Asn Arg Leu Leu Ala Leu Ala Met  
 245 250 255  
 Val Thr Val Val Trp Leu Thr Leu Glu Ser Val Val Thr Arg Val Ala  
 260 265 270  
 Val Leu Val Val Leu Leu Cys Leu Ala Pro Val Tyr Ala Ser Arg Cys  
 275 280 285  
 Thr His Leu Glu Asn Arg Asp Phe Val Thr Gly Thr Gln Gly Thr Thr  
 290 295 300  
 Arg Val Thr Leu Val Leu Glu Leu Gly Gly Cys Val Thr Ile Thr Ala  
 305 310 315 320  
 Glu Gly Lys Pro Ser Met Asp Val Trp Leu Asp Ala Ile Tyr Gln Glu  
 325 330 335  
 Asn Pro Ala Gln Thr Arg Glu Tyr Cys Leu His Ala Lys Leu Ser Asp  
 340 345 350  
 Thr Lys Val Ala Ala Arg Cys Pro Thr Met Gly Pro Ala Thr Leu Ala  
 355 360 365  
 Glu Glu His Gln Gly Gly Thr Val Cys Lys Arg Asp Gln Ser Asp Arg  
 370 375 380  
 Gly Trp Gly Asn His Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala  
 385 390 395 400  
 Cys Val Lys Ala Ala Cys Glu Ala Lys Lys Lys Ala Thr Gly His Val  
 405 410 415  
 Tyr Asp Ala Asn Lys Ile Val Tyr Thr Val Lys Glu Pro His Thr Gly  
 420 425 430  
 Asp Tyr Val Ala Ala Asn Glu Thr His Ser Gly Arg Lys Thr Ala Ser  
 435 440 445  
 Phe Thr Val Ser Ser Glu Lys Thr Ile Leu Thr Met Gly Glu Tyr Gly  
 450 455 460  
 Asp Val Ser Leu Leu Cys Arg Val Ala Ser Gly Val Asp Leu Ala Gln  
 465 470 475 480  
 Thr Val Ile Leu Glu Leu Asp Lys Thr Val Glu His Leu Pro Thr Ala  
 485 490 495  
 Trp Gln Val His Arg Asp Trp Phe Asn Asp Leu Ala Leu Pro Trp Lys  
 500 505 510  
 His Glu Gly Ala Arg Asn Trp Asn Asn Ala Glu Arg Leu Val Glu Phe  
 515 520 525  
 Gly Ala Pro His Ala Val Lys Met Asp Val Tyr Asn Leu Gly Asp Gln

-continued

| 530 |     |     | 535 |     |     | 540 |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Gly | Val | Leu | Leu | Lys | Ala | Leu | Ala | Gly | Val | Pro | Val | Ala | His | Ile |
| 545 |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     |     | 560 |
| Glu | Gly | Thr | Lys | Tyr | His | Leu | Lys | Ser | Gly | His | Val | Thr | Cys | Glu | Val |
|     |     |     | 565 |     |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Gly | Leu | Glu | Lys | Leu | Lys | Met | Lys | Gly | Leu | Thr | Tyr | Thr | Met | Cys | Asp |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     | 590 |     |     |     |
| Lys | Thr | Lys | Phe | Thr | Trp | Lys | Arg | Ala | Pro | Thr | Asp | Ser | Gly | His | Asp |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |
| Thr | Val | Val | Met | Glu | Val | Thr | Phe | Ser | Gly | Thr | Lys | Pro | Cys | Arg | Ile |
| 610 |     |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |
| Pro | Val | Arg | Ala | Val | Ala | His | Gly | Ser | Pro | Asp | Val | Asn | Val | Ala | Met |
| 625 |     |     |     | 630 |     |     |     |     |     | 635 |     |     |     |     | 640 |
| Leu | Ile | Thr | Pro | Asn | Pro | Thr | Ile | Glu | Asn | Asn | Gly | Gly | Gly | Phe | Ile |
|     |     |     | 645 |     |     |     |     |     | 650 |     |     |     |     | 655 |     |
| Glu | Met | Gln | Leu | Pro | Pro | Gly | Asp | Asn | Ile | Ile | Tyr | Val | Gly | Glu | Leu |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |
| Ser | Tyr | Gln | Trp | Phe | Gln | Lys | Gly | Ser | Ser | Ile | Gly | Arg | Val | Phe | Gln |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |
| Lys | Thr | Lys | Lys | Gly | Ile | Glu | Arg | Leu | Thr | Val | Ile | Gly | Glu | His | Ala |
|     | 690 |     |     |     | 695 |     |     |     |     |     | 700 |     |     |     |     |
| Trp | Asp | Phe | Gly | Ser | Ala | Gly | Gly | Phe | Leu | Ser | Ser | Ile | Gly | Lys | Ala |
| 705 |     |     |     | 710 |     |     |     |     |     | 715 |     |     |     |     | 720 |
| Leu | His | Thr | Val | Leu | Gly | Gly | Ala | Phe | Asn | Ser | Ile | Phe | Gly | Gly | Val |
|     |     |     | 725 |     |     |     |     |     | 730 |     |     |     |     | 735 |     |
| Gly | Phe | Leu | Pro | Lys | Leu | Leu | Leu | Gly | Val | Ala | Leu | Ala | Trp | Leu | Gly |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |
| Leu | Asn | Met | Arg | Asn | Pro | Thr | Met | Ser | Met | Ser | Phe | Leu | Leu | Ala | Gly |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |
| Val | Leu | Val | Leu | Ala | Met | Thr | Leu | Gly | Val | Gly | Ala | Asp | Gln | Gly | Cys |
|     | 770 |     |     |     | 775 |     |     |     |     |     | 780 |     |     |     |     |
| Ala | Ile | Asn | Phe | Gly | Lys | Arg | Glu | Leu |     |     |     |     |     |     |     |
| 785 |     |     |     |     | 790 |     |     |     |     |     |     |     |     |     |     |

<210> SEQ ID NO 44  
 <211> LENGTH: 2500  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 44

```

agtaaatcct gtgtgctaatt tgagggtgcat tggctctgcaa atcgagttgc taggcaataa 60
acacatttgg attaattttaa atcgttcgtt gagcgattag cagagaactg accagaacat 120
gtctggctcgt aaagctcagg gaaaaaacct gggcgtcaat atggtacgac gaggagtctg 180
ctccttgta aacaaaataa aacaaaaaac aaaacaaatt ggaaacagac ctggaccttc 240
aagaggtgtt caaggattta tctttttctt tttgttcaac attttgactg gaaaaaagat 300
cacagcccac ctaagaggt tgtggaaaat gctggacca agacaaggct tggctgttct 360
aaggaaagtc aagagagtgg tggccagttt gatgagagga ttgtcctcaa ggaaacgccc 420
ttcccatgat gttctgactg tgcaattcct aattttgggc atgctggctt gtgtcggagc 480
agctaccgtg cgaagaagac gcgacggaag caccgtgata agggctgagg gtaaggatgc 540
ggctacgcag gtgagagtag agaatggcac ttgcgtaata ctgcgactg atatgggatc 600

```

-continued

---

```

ctggtgtgac gatagcctca gttatgaatg cgtaacaata gaccagggcg aagaacctgt 660
ggacgttgac tgtttctgta gaaatgtgga tggcgtttat ctggagtacg gccgctgtgg 720
aaaacaggag ggctccacgaa ctcgaagatc tgtgctgatt ccaagtcacg cgcaaggaga 780
gttgaccggg agaggccaca agtggcttga aggggactca ttgaggaccc acctgactag 840
ggtggagggt tgggtttgga agaatcgggt gctcgcgctc gctatggtea ccgtcgtgtg 900
gctgacactg gagagtgtcg tgactcgggt tgctgtgttg gttgtcctcc tctgtttggc 960
cccagtgtac gcgtccaggt gtactcattt ggaaaacaga gatthttgtca ccggcaccca 1020
ggggacgact cgggtaaccc tgggtcttga actgggtggg tgcgttacta ttaccgctga 1080
gggcaaaccc tctatggatg tgtggctgga tgcaatctat caggagaatc ccgcacaaac 1140
caggaatat tgcttcacg caaagctgtc cgatacaaaag gtcgcggcta ggtgcccaac 1200
aatgggacgg gccaccctgg cggaggaaca tcaggagggt acagtgtgca aacgggacca 1260
gagtgataga ggctggggta atcactgctg cctgttcggc aaaggaagta ttgtcgttg 1320
cgtcaaggca gcctgtgagg ccaaaaagaa ggctactggg cacgtctatg acgccaacaa 1380
gatcgtttat acagtgaag tgaaccaca cacaggggat tacgtggcgg ccaacgagac 1440
tcattccggg cgcaaacgg ccagcttcac cgtgtcatcc gaaaagacca tcctcactat 1500
gggggagtat ggcgacgttt ctctgctctg ccgggtgggt agcggagtcg acctggccca 1560
gacagtcacg ctggaactgg ataaaacagt tgagcatctg cctaccgctt ggcaggtgca 1620
cagggattgg tttaacgacc ttcccctgcc atggaacat gaaggagcga gaaactggaa 1680
taatgcagag cgactcgtag aattcgggtc ccctcatgcc gtgaagatgg acgtctacaa 1740
tctgggtgat cagaccggcg ttctccttaa agctctcgtt ggcgtaccag ttgccacat 1800
cgaaggaacg aagtaccacc tgaagtcagg ccatgtaact tgcgaggtgg gcctggagaa 1860
gtgaaaaatg aaaggtctta cgtacacaat gttgacaag accaagttca catggaagag 1920
ggccccaca gatagcggcc acgatactgt ggtgatggag gtgacctttt ctggaacaaa 1980
acctgcaga ataccctgct gggctgtgac tcacggatct cccgatgtca atgttctat 2040
gctgattaca cctaacccta ccatcgagaa taacgggtgg ggthttattg agatgcagct 2100
tccgccaggc gataacatca tctacgtggg cgaactctct taccagtggg ttcagaaagg 2160
gagttcaatt gggcgggtct tccaaaaaac gaagaagga atcgaaacgat tgacggttat 2220
cggcgagcac gcattgggatt ttggttcgc agggggattc ctgtcttcta ttggtaaggc 2280
actgcatacc gtgctggggg gcgcattcaa ttctatthtc gggggcgtgg ggttcctgcc 2340
taaaactctg ctgggagtag cctggcctg gttgggactg aatatgcgga atccgacgat 2400
gtccatgtca ttctcttgg ccggcgtgct tgtaactggc atgacactgg gcgttggcgc 2460
cgatcaagga tgcgccatca actttggcaa gagagagctc 2500

```

&lt;210&gt; SEQ ID NO 45

&lt;211&gt; LENGTH: 2500

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 45

```

tcatttagga cacacgatta actccacgta accagacggt tagctcaacg atccgttatt 60
tgtgtaaacc taattaaat tagcaagcaa ctcgctaate gtctcttgac tggctcttga 120

```

-continued

---

|                                                                     |      |
|---------------------------------------------------------------------|------|
| cagaccagca tttcagtgcc ctttttggga cccgcagtta taccatgctg ctcccaage    | 180  |
| gaggaacagt ttgttttatt ttgttttttg ttttgtttaa cctttgtctg gacctggaag   | 240  |
| ttctccacaa gttcctaaat agaaaaagaa aaacaagttg taaaactgac cttttttcta   | 300  |
| gtgtcgggtg gatttctoca acaccttta cgacctgggt tctgttccga accgacaaga    | 360  |
| ttcctttcag ttctctcacc accggtaaaa ctactctcct aacaggagtt cctttgcggc   | 420  |
| aagggtacta caagactgac acgttaagga ttaaaaccgg tacgaccgaa cacagcctcg   | 480  |
| tcgatggcac gcttttcttg cgctgccttc gtggcaactat tcccgaactcc cattcctacg | 540  |
| ccgatgcgtc cactctcctc tcttaccttg aacgcattat gagcgtgac tataccttag    | 600  |
| gaccacactg ctatcggagt caatacttac gcattgttat ctgggtccgc ttcttggaaca  | 660  |
| cctgcaactg acaaagacat ctttacacct accgcaaata gacctcatgc cggcgacacc   | 720  |
| ttttgtcttc ccgagtgtct gagcttctag acacgactaa gggtcagtc gcgttctctc    | 780  |
| caactggcca tctccggtgt tcaccgaact tcccctgagt aactcctggg tggactgatc   | 840  |
| ccacctcca acccaaacct tcttagccaa cgagcgcgag cgataccagt ggcagcacac    | 900  |
| cgactgtgac ctctcacagc actgagccca acgacacaac caacaggagg agacaaaaccg  | 960  |
| gggtcacatg cgcaggtcca catgagtaaa ccttttgtct ctaaaacagt ggcctgggt    | 1020 |
| cccctgctga gcccatggg accacgaact tgaccacca acgcaatgat aatggcgact     | 1080 |
| cccgtttggg agatacctac acaccgacct acgttagata gtccctcttag ggcgtgtttg  | 1140 |
| gtcccttata acggaagtgc gtttcgacag getatgttcc cagcgcgat ccacgggttg    | 1200 |
| ttaccctggc cgggtggacc gcctccttgt agtccctcca tgtcacagct ttgcctgggt   | 1260 |
| ctcactatct ccgaccccat tagtgacgcc ggacaagccg tttccttcat aacagcgaac   | 1320 |
| gcagtccgt cggacactcc ggtttttctt ccgatgacc gtgcagatac tgcggtgtgt     | 1380 |
| ctagcaaata tgtcacttcc accttgggtg gtgtccccta atgcaccgcc ggttgcctcg   | 1440 |
| agtaaggcca gcgttttgcc ggtcgaagtg gcacagtagg cttttctggg aggagtgata   | 1500 |
| ccccctcata ccgctgcaaa gagacgagac ggcccaccga tcgcctcagc tggaccgggt   | 1560 |
| ctgtcagtag gaccttgacc tattttgtca actcgtagac ggatggcgaa ccgtccactg   | 1620 |
| gtccctaacc aaattgtctg aacgggacgg tacctttgta ctccctcgt ctttgacctt    | 1680 |
| attacgtctc gctgagcctc ttaagccacg gggagtacgg cacttctacc tgcagatgtt   | 1740 |
| agaccacta gtctggcgc aagaggaatt tcgagagcga ccgcatggc aacgggtgta      | 1800 |
| gcttccttgc ttcattgttg acttcagtc ggtacattga acgctccacc cggacctctt    | 1860 |
| caacttttac tttccagaat gcatgtgta cacactgttc tggttcaagt gtaccttctc    | 1920 |
| ccgggggtgt ctatcgcgg tgctatgaca ccactacctc cactggaaaa gacctgttt     | 1980 |
| tgggacgtct tatgggacg cccgacatcg agtgcctaga gggctacagt tacaacgata    | 2040 |
| cgactaatgt ggattgggat ggtagctctt attgccacca ccaaaataac tctacgtcga   | 2100 |
| aggcggctcg ctattgtagt agatgcacc gcttgagaga atggtcacca aagtctttcc    | 2160 |
| ctcaagttaa cccgccaga aggttttttg cttcttccct tagcttgcta actgccaata    | 2220 |
| gccgctcgtg cgtaccctaa aaccaaggcg tcccctaag gacagaagat aaccattccg    | 2280 |
| tgacgtatgg cacgaccccc cgcgtaagtt aagataaaag cccccgacc ccaaggacgg    | 2340 |
| atgtgaggac gacctctcgc gggaccggac caacctgac ttatacgcct taggctgcta    | 2400 |
| caggtagcgt aaggagaacc ggccgcacga acatgaccgg tactgtgacc cgcaaccgcg   | 2460 |
| gctagttcct acgcgtagt tgaaacggtt ctctctcgag                          | 2500 |

-continued

---

```

<210> SEQ ID NO 46
<211> LENGTH: 794
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 46
Met Ser Gly Arg Lys Ala Gln Gly Lys Thr Leu Gly Val Asn Met Val
1 5 10 15
Arg Arg Gly Val Arg Ser Leu Ser Asn Lys Ile Lys Gln Lys Thr Lys
20 25 30
Gln Ile Gly Asn Arg Pro Gly Pro Ser Arg Gly Val Gln Gly Phe Ile
35 40 45
Phe Phe Phe Leu Phe Asn Ile Leu Thr Gly Lys Lys Ile Thr Ala His
50 55 60
Leu Lys Arg Leu Trp Lys Met Leu Asp Pro Arg Gln Gly Leu Ala Val
65 70 75 80
Leu Arg Lys Val Lys Arg Val Val Ala Ser Leu Met Arg Gly Leu Ser
85 90 95
Ser Arg Lys Arg Arg Ser His Asp Val Leu Thr Val Gln Phe Leu Ile
100 105 110
Leu Gly Met Leu Gly Met Thr Ile Ala Ala Thr Val Arg Arg Glu Arg
115 120 125
Asp Gly Ser Met Val Ile Arg Ala Glu Gly Arg Asp Ala Ala Thr Gln
130 135 140
Val Arg Val Glu Asn Gly Thr Cys Val Ile Leu Ala Thr Asp Met Gly
145 150 155 160
Ser Trp Cys Asp Asp Ser Leu Ala Tyr Glu Cys Val Thr Ile Asp Gln
165 170 175
Gly Glu Glu Pro Val Asp Val Asp Cys Phe Cys Arg Gly Val Glu Lys
180 185 190
Val Thr Leu Glu Tyr Gly Arg Cys Gly Arg Arg Glu Gly Ser Arg Ser
195 200 205
Arg Arg Ser Val Leu Ile Pro Ser His Ala Gln Arg Asp Leu Thr Gly
210 215 220
Arg Gly His Gln Trp Leu Glu Gly Glu Ala Val Lys Ala His Leu Thr
225 230 235 240
Arg Val Glu Gly Trp Val Trp Lys Asn Lys Leu Phe Thr Leu Ser Leu
245 250 255
Val Met Val Ala Trp Leu Met Val Asp Gly Leu Leu Pro Arg Ile Leu
260 265 270
Ile Val Val Val Ala Leu Ala Leu Ala Pro Ala Tyr Ala Ser Arg Cys
275 280 285
Thr His Leu Glu Asn Arg Asp Phe Val Thr Gly Val Gln Gly Thr Thr
290 295 300
Arg Leu Thr Leu Val Leu Glu Leu Gly Gly Cys Val Thr Val Thr Ala
305 310 315 320
Asp Gly Lys Pro Ser Leu Asp Val Trp Leu Asp Ser Ile Tyr Gln Glu
325 330 335
Ser Pro Ala Gln Thr Arg Glu Tyr Cys Leu His Ala Lys Leu Thr Gly
340 345 350
Thr Lys Val Ala Ala Arg Cys Pro Thr Met Gly Pro Ala Thr Leu Pro
355 360 365

```

-continued

---

Glu Glu His Gln Ser Gly Thr Val Cys Lys Arg Asp Gln Ser Asp Arg  
 370 375 380

Gly Trp Gly Asn His Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Thr  
 385 390 395 400

Cys Val Lys Val Thr Cys Glu Asp Lys Lys Lys Ala Thr Gly His Val  
 405 410 415

Tyr Asp Val Asn Lys Ile Thr Tyr Thr Ile Lys Val Glu Pro His Thr  
 420 425 430

Gly Glu Phe Val Ala Ala Asn Glu Thr His Ser Gly Arg Lys Ser Ala  
 435 440 445

Ser Phe Thr Val Ser Ser Glu Lys Thr Ile Leu Thr Leu Gly Asp Tyr  
 450 455 460

Gly Asp Val Ser Leu Leu Cys Arg Val Ala Ser Gly Val Asp Leu Ala  
 465 470 475 480

Gln Thr Val Val Leu Ala Leu Asp Lys Thr His Glu His Leu Pro Thr  
 485 490 495

Ala Trp Gln Val His Arg Asp Trp Phe Asn Asp Leu Ala Leu Pro Trp  
 500 505 510

Lys His Asp Gly Ala Glu Ala Trp Asn Glu Ala Gly Arg Leu Val Glu  
 515 520 525

Phe Gly Thr Pro His Ala Val Lys Met Asp Val Phe Asn Leu Gly Asp  
 530 535 540

Gln Thr Gly Val Leu Leu Lys Ser Leu Ala Gly Val Pro Val Ala Ser  
 545 550 555 560

Ile Glu Gly Thr Lys Tyr His Leu Lys Ser Gly His Val Thr Cys Glu  
 565 570 575

Val Gly Leu Glu Lys Leu Lys Met Lys Gly Leu Thr Tyr Thr Val Cys  
 580 585 590

Asp Lys Thr Lys Phe Thr Trp Lys Arg Ala Pro Thr Asp Ser Gly His  
 595 600 605

Asp Thr Val Val Met Glu Val Gly Phe Ser Gly Thr Arg Pro Cys Arg  
 610 615 620

Ile Pro Val Arg Ala Val Ala His Gly Val Pro Glu Val Asn Val Ala  
 625 630 635 640

Met Leu Ile Thr Pro Asn Pro Thr Met Glu Asn Asn Gly Gly Gly Phe  
 645 650 655

Ile Glu Met Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr Val Gly Asp  
 660 665 670

Leu Asp His Gln Trp Phe Gln Lys Gly Ser Ser Ile Gly Arg Val Leu  
 675 680 685

Gln Lys Thr Arg Lys Gly Ile Glu Arg Leu Thr Val Leu Gly Glu His  
 690 695 700

Ala Trp Asp Phe Gly Ser Val Gly Gly Val Met Thr Ser Ile Gly Arg  
 705 710 715 720

Ala Met His Thr Val Leu Gly Gly Ala Phe Asn Thr Leu Leu Gly Gly  
 725 730 735

Val Gly Phe Leu Pro Lys Ile Leu Leu Gly Val Ala Met Ala Trp Leu  
 740 745 750

Gly Leu Asn Met Arg Asn Pro Thr Leu Ser Met Gly Phe Leu Leu Ser  
 755 760 765

Gly Gly Leu Val Leu Ala Met Thr Leu Gly Val Gly Ala Asp Gln Gly  
 770 775 780

-continued

Cys Ala Ile Asn Phe Gly Lys Arg Glu Leu  
785 790

<210> SEQ ID NO 47  
<211> LENGTH: 2500  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 47

```

agtaaatcct gtgtgctaata tgaggatgcat tgggtctgcaa atcgagtgc taggcaataa 60
acacatttgg attaatttta atcgttcgtt gagcgattag cagagaactg accagaacat 120
gtctggctcg aaagctcagg gaaaaaccct gggcgtcaat atggtacgac gaggagtctg 180
ctccttgta aacaaaataa aacaaaaaac aaaacaaatt ggaacagac ctggacctc 240
aagaggtgtt caaggattta tcttttctt tttgtcaac atttgactg gaaaaagat 300
cacagcccac ctaagaggt tgtggaat gctggacca agacaaggct tggctgttct 360
aaggaaagtc aagagagtgg tggccagtt gatgagagga ttgtcctca ggaacgccc 420
ttcccatgat gttctgactg tgcaattcct aatttgggc atgctgggga tgacgatcgc 480
agctactgtg cgaagggaga gagacggctc tatggtgatc agagccgaag gtagggacgc 540
tgcgaccag gtgaggtgctg aaaatggcac ctgtgttatt ctggcgaccg acatgggctc 600
ctggtgtgat gattctctgg cttatgaatg tgttactatt gatcagggtg aagagcctgt 660
ggacgtggac tgtttctgta gaggcgtcga gaaagtgacc ctggaatatg gacgatgtgg 720
cggcgagaa ggtccagga gtcggagatc cgtgttgatc ccttcacatg cgcagcgcga 780
tctgacaggg aggggtcacc agtggctcga aggcgaagca gtcaaggccc atctgactcg 840
cgttgaaggg tgggtgtgga aaaacaaact cttaccctt agcctggtga tggctcgcgtg 900
gctgatggta gacggactcc tccccgcat tctcattgtt gtggtggctc tcgctcgcgc 960
ccctgcatac gcgtccaggt gtacgcacct cgaaaatcga gatttcgtca caggcgtcca 1020
aggtaactacc cggctcacc cctgtctgga gctgggagggc tgtgtcactg ttacagccga 1080
cggaaaacct agtctggatg tgtggctgga ctccatctat caggagagcc cggcacagac 1140
cagggagtac tgcctccacg ctaagctgac tgggacaaag gtagccgcaa gatgtcccac 1200
aatggggcct gccacctgac ccgaggaaca ccaatccggt acggtatgca agcgagatca 1260
gtctgatcgc ggtggggga atcattgagg cctcttcggt aaaggcagca ttgtcacttg 1320
cgtgaaggtg acatgcgagg acaagaagaa ggccacaggt catgtatatg atgtgaacaa 1380
aatcacatat accattaagg tagaaccaca tacaggggaa ttcgtggcag caaacgagac 1440
tcatagcggg cgaagtcgg cctccttcac cgtctcctcc gagaaaaaca tctgaccct 1500
cgggagctac ggcgacgtat ctttctgtg cagggtggcc agcggcgtgg accttgetca 1560
gacagtctgt ttggccctgg acaagacaca tgagcacttg ccaacagcct ggcaggtgca 1620
cagggactgg ttaacgacc tggcctccc gtggaacat gacggcgtg aagcatggaa 1680
tgaggcaggg agactggtgg aatttggaa cccacacgcc gtaaagatgg acgttttcaa 1740
tcttgggtgac cagacagggg tgctcctgaa atcactggcg ggcgtgcctg tagccagcat 1800
cgagggcaca aagtatcacc tgaagtctgg gcatgtaacc tgcaagtgg gcctgaaaa 1860
gctgaagatg aaaggactta cgtacactgt ttgtgataag accaagtta catggaagcg 1920
agccccaacg gattccggcc atgataccgt cgtgatggag gttggtttct ccggcaccag 1980

```

-continued

---

```

accatgtaga ataccagtga gagctgtcgc ccacggtgta cccgaggtaa acgtggccat 2040
gctgattaca ccgaatccca ctatggagaa caatggcgga gggttcatcg aaatgcagct 2100
gccgcctgga gacaacatca tttatgtcgg cgacctgat catcaatggt tccagaaagg 2160
gtcttccatc ggccgcgtcc ttcagaagac acgaaaaggc attgaaagac ttacagtcc 2220
ggcgaaacat gcctgggact tcgggtcagt tggcggggta atgacaagca taggcagagc 2280
tatgcacacc gttctcggtg gggcatttaa tactctgttg ggtggcgtgg gttttcttcc 2340
gaaaatcctg ctcggtgtcg caatggcctg gcttggactg aatatgcgca atcctacact 2400
gagtatgggg tttcttctgt caggaggcct ggtcctggca atgactctgg gagtgggcg 2460
cgatcaagga tgcgccatca actttggcaa gagagagctc 2500

```

```

<210> SEQ ID NO 48
<211> LENGTH: 2500
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 48

```

```

tcatttagga cacacgatta actccacgta accagacggt tagctcaacg atccgttatt 60
tgtgtaaacc taattaaaat tagcaagcaa ctcgctaacc gtctcttgac tggctctgta 120
cagaccagca tttcagagtc ctttttggga cccgcagtta taccatgctg ctctcaagc 180
gaggaaacagt ttgttttatt ttgttttttg ttttgtttaa cctttgtctg gacctggaag 240
ttctccacaa gttcctaaat agaaaaagaa aaacaagttg taaaactgac cttttttcta 300
gtgtcgggtg gatttctoca acaccttta cgacctgggt tctgttccga accgacaaga 360
ttcctttcag ttctctcacc accggtcaaa ctactctcct aacaggaggt cctttgoggc 420
aagggtacta caagactgac acgttaagga ttaaaaccog tacgaccctt actgctagcg 480
tcgatgacac gcttccctct ctctgcogag ataccactag tctcggett ccatccctg 540
acgctgggtc cactcccagc ttttacogtg gacacaataa gaccgctggc tgtaccogag 600
gaccacacta ctaagagacc gaatacttac acaatgataa ctagtcccac ttctcggaca 660
cctgcacctg acaaaagacat ctccgcagct ctttactgg gaccttatac ctgctacacc 720
ggccgctctt ccgaggtcct cagcctctag gcacaactag ggaagtgtac gcgtcgcgct 780
agactgtccc tccccagtgg tcaccgagct tccgcttctg cagttccggg tagactgagc 840
gcaacttccg acccacacct ttttgtttga gaaatgggaa tcggaccact accagcgcac 900
cgactaccat ctgctgaggg aaggggogta agagtaacaa caccaccgag agcgcgagcg 960
gggacgtatg cgcaggtcca catgctgga gcttttagct ctaaagcagt gtcgcgaggt 1020
tccatgatgg gccagatggg agcaccacct cgacctccg acacagtgac aatgtcggct 1080
gccttttggg tcagacctac acaccgacct gaggtagata gtctctcgg gccgtgtctg 1140
gtccctcatg acggaggtgc gattcagctg accctgttcc catcggcggt ctacagggtg 1200
ttaccccgga cgggtgaaac ggctccttgt ggttaggcca tgccatacgt tcgctctagt 1260
cagactagcg cctacccctc tagtaacgcc ggagaagcca tttcogtctg aacagtgaac 1320
gcacttccac tgtacgctcc tgttcttctt ccggtgtcca gtacatatac tacacttgtt 1380
ttagtgtata tggtaattcc atcttgggtg atgtccctt aagcaccgct gtttgetctg 1440
agtatcgctt gctttcaggc ggaggaagtg gcagaggagg ctcttttgtt aggactggga 1500
gcctctgatg ccgctgcata gaaacgacac gtcccaccgg tcgcccaccc tggaaacgagt 1560

```

-continued

---

```

ctgtcagcac aaccgggacc tgttctgtgt actcgtgaac gggtgtcgga cegtccacgt 1620
gtccctgacc aaattgctgg accgcgaggg cacctttgta ctgcccgcac ttcgtacctt 1680
actccgtccc tctgaccacc ttaaaccttg ggggtgtcgg catttctacc tgcaaaagt 1740
agaaccactg gtctgtcccc acgaggactt tagtgaccgc ccgcacggac atcggtcgta 1800
gctcccgtgt ttcatagtgg acttcagacc cgtacattgg acgcttcacc cggacctttt 1860
cgacttttac tttctgaat gcatgtgaca aacactattc tggttcaaat gtaccttcgc 1920
tcgggggtgc ctaaggccgg tactatggca gcaactacc ccaacaaaga ggccgtggtc 1980
tggtagacatc tatggtcact ctcgacagcg ggtgccacat gggctccatt tgcaccggta 2040
cgactaatgt ggcttagggt gatacctctt gttaccgcct cccaagtagc tttacgtcga 2100
cggcggacct ctgtttagt aaatacagcc gctggageta gtagttacca aggtctttcc 2160
cagaaggtag ccggcgcagg aagtctctg tgcctttccg taactttctg aatgtcagga 2220
cccgtttgta cggacctga agcccagtca accgccccat tactgttcgt atcgtctcg 2280
atacgtgtgg caagagccac ccgtaaat atgagacaac ccaccgcacc caaaagaagg 2340
cttttaggac gagccacagc gttaccggac cgaacctgac ttatacgcgt taggatgtga 2400
ctcatacccc aaagaagaca gtctctcgga ccaggaccgt tactgagacc ctcaccgcgc 2460
gctagttcct acgcggtagt tgaaacggtt ctctctcgag 2500

```

```

<210> SEQ ID NO 49
<211> LENGTH: 791
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 49

```

```

Met Ser Gly Arg Lys Ala Gln Gly Lys Thr Leu Gly Val Asn Met Val
1 5 10 15
Arg Arg Gly Val Arg Ser Leu Ser Asn Lys Ile Lys Gln Lys Thr Lys
20 25 30
Gln Ile Gly Asn Arg Pro Gly Gly Val Gln Gly Phe Ile Phe Phe Phe
35 40 45
Leu Phe Asn Ile Leu Thr Gly Lys Lys Ile Thr Ala His Leu Lys Arg
50 55 60
Leu Trp Lys Met Leu Asp Pro Arg Gln Gly Leu Ala Val Leu Arg Lys
65 70 75 80
Val Lys Arg Val Val Ala Ser Leu Met Arg Gly Leu Ser Ser Arg Lys
85 90 95
Arg Arg Ser His Asp Val Leu Thr Val Gln Phe Leu Ile Leu Gly Met
100 105 110
Leu Gly Met Thr Ile Ala Ala Thr Val Arg Lys Glu Arg Asp Gly Ser
115 120 125
Thr Val Ile Arg Ala Glu Gly Lys Asp Ala Ala Thr Gln Val Arg Val
130 135 140
Glu Asn Gly Thr Cys Val Ile Leu Ala Thr Asp Met Gly Ser Trp Cys
145 150 155 160
Asp Asp Ser Leu Ser Tyr Glu Cys Val Thr Ile Asp Gln Gly Glu Glu
165 170 175
Pro Val Asp Val Asp Cys Phe Cys Arg Asn Val Asp Gly Val Tyr Leu
180 185 190

```

-continued

---

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Tyr | Gly | Arg | Cys | Gly | Lys | Gln | Glu | Gly | Ser | Arg | Thr | Arg | Arg | Ser |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |
| Val | Leu | Ile | Pro | Ser | His | Ala | Gln | Gly | Glu | Leu | Thr | Gly | Arg | Gly | His |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |
| Lys | Trp | Leu | Glu | Gly | Asp | Ser | Leu | Arg | Thr | His | Leu | Thr | Arg | Val | Glu |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |
| Gly | Trp | Val | Trp | Lys | Asn | Arg | Leu | Leu | Ala | Leu | Ala | Met | Val | Thr | Val |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |
| Val | Trp | Leu | Thr | Leu | Glu | Ser | Val | Val | Thr | Arg | Val | Ala | Val | Leu | Val |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |
| Val | Leu | Leu | Cys | Leu | Ala | Pro | Val | Tyr | Ala | Ser | Arg | Cys | Thr | His | Leu |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |
| Glu | Asn | Arg | Asp | Phe | Val | Thr | Gly | Thr | Gln | Gly | Thr | Thr | Arg | Val | Thr |
|     | 290 |     |     |     |     | 295 |     |     |     |     |     | 300 |     |     |     |
| Leu | Val | Leu | Glu | Leu | Gly | Gly | Cys | Val | Thr | Ile | Thr | Ala | Glu | Gly | Lys |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |
| Pro | Ser | Met | Asp | Val | Trp | Leu | Asp | Ala | Ile | Tyr | Gln | Glu | Asn | Pro | Ala |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |
| Gln | Thr | Arg | Glu | Tyr | Cys | Leu | His | Ala | Lys | Leu | Ser | Asp | Thr | Lys | Val |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     |     | 350 |     |
| Ala | Ala | Arg | Cys | Pro | Thr | Met | Gly | Pro | Ala | Thr | Leu | Ala | Glu | Glu | His |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
| Gln | Gly | Gly | Thr | Val | Cys | Lys | Arg | Asp | Gln | Ser | Asp | Arg | Gly | Trp | Gly |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Asn | His | Cys | Gly | Leu | Phe | Gly | Lys | Gly | Ser | Ile | Val | Ala | Cys | Val | Lys |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |
| Ala | Ala | Cys | Glu | Ala | Lys | Lys | Lys | Ala | Thr | Gly | His | Val | Tyr | Asp | Ala |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Asn | Lys | Ile | Val | Tyr | Thr | Val | Lys | Val | Glu | Pro | His | Thr | Gly | Asp | Tyr |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |
| Val | Ala | Ala | Asn | Glu | Thr | His | Ser | Gly | Arg | Lys | Thr | Ala | Ser | Phe | Thr |
|     |     |     | 435 |     |     |     | 440 |     |     |     |     | 445 |     |     |     |
| Val | Ser | Ser | Glu | Lys | Thr | Ile | Leu | Thr | Met | Gly | Glu | Tyr | Gly | Asp | Val |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Ser | Leu | Leu | Cys | Arg | Val | Ala | Ser | Gly | Val | Asp | Leu | Ala | Gln | Thr | Val |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |
| Ile | Leu | Glu | Leu | Asp | Lys | Thr | Val | Glu | His | Leu | Pro | Thr | Ala | Trp | Gln |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |
| Val | His | Arg | Asp | Trp | Phe | Asn | Asp | Leu | Ala | Leu | Pro | Trp | Lys | His | Glu |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |
| Gly | Ala | Arg | Asn | Trp | Asn | Asn | Ala | Glu | Arg | Leu | Val | Glu | Phe | Gly | Ala |
|     |     |     | 515 |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Pro | His | Ala | Val | Lys | Met | Asp | Val | Tyr | Asn | Leu | Gly | Asp | Gln | Thr | Gly |
|     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Val | Leu | Leu | Lys | Ala | Leu | Ala | Gly | Val | Pro | Val | Ala | His | Ile | Glu | Gly |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Thr | Lys | Tyr | His | Leu | Lys | Ser | Gly | His | Val | Thr | Cys | Glu | Val | Gly | Leu |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Glu | Lys | Leu | Lys | Met | Lys | Gly | Leu | Thr | Tyr | Thr | Met | Cys | Asp | Lys | Thr |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Lys | Phe | Thr | Trp | Lys | Arg | Ala | Pro | Thr | Asp | Ser | Gly | His | Asp | Thr | Val |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |
| Val | Met | Glu | Val | Thr | Phe | Ser | Gly | Thr | Lys | Pro | Cys | Arg | Ile | Pro | Val |

-continued

| 610                                                                                                                                         | 615 | 620 |
|---------------------------------------------------------------------------------------------------------------------------------------------|-----|-----|
| Arg Ala Val Ala His Gly Ser Pro Asp Val Asn Val Ala Met Leu Ile<br>625                    630                    635                    640 |     |     |
| Thr Pro Asn Pro Thr Ile Glu Asn Asn Gly Gly Gly Phe Ile Glu Met<br>645                    650                    655                        |     |     |
| Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr Val Gly Glu Leu Ser Tyr<br>660                    665                    670                        |     |     |
| Gln Trp Phe Gln Lys Gly Ser Ser Ile Gly Arg Val Phe Gln Lys Thr<br>675                    680                    685                        |     |     |
| Lys Lys Gly Ile Glu Arg Leu Thr Val Ile Gly Glu His Ala Trp Asp<br>690                    695                    700                        |     |     |
| Phe Gly Ser Ala Gly Gly Phe Leu Ser Ser Ile Gly Lys Ala Leu His<br>705                    710                    715                    720 |     |     |
| Thr Val Leu Gly Gly Ala Phe Asn Ser Ile Phe Gly Gly Val Gly Phe<br>725                    730                    735                        |     |     |
| Leu Pro Lys Leu Leu Leu Gly Val Ala Leu Ala Trp Leu Gly Leu Asn<br>740                    745                    750                        |     |     |
| Met Arg Asn Pro Thr Met Ser Met Ser Phe Leu Leu Ala Gly Val Leu<br>755                    760                    765                        |     |     |
| Val Leu Ala Met Thr Leu Gly Val Gly Ala Asp Gln Gly Cys Ala Ile<br>770                    775                    780                        |     |     |
| Asn Phe Gly Lys Arg Glu Leu<br>785                    790                                                                                   |     |     |

&lt;210&gt; SEQ ID NO 50

&lt;211&gt; LENGTH: 2491

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 50

```

agtaaatcct gtgtgctaata tgagggtgcat tggctctgcaa atcgagttgc taggcaataa 60
acacatttgg attaatatta atcggtcggt gagcgattag cagagaactg accagaacat 120
gtctggctcg aaagctcagg gaaaaaacct gggcgtaaat atggtacgac gaggagtctg 180
ctccttgtca aacaaaataa aacaaaaaac aaaacaaatt gaaacagac ctggaggtgt 240
tcaaggattt atctttttct ttttgttcaa cattttgact ggaaaaaaga tcacagccca 300
cctaaagagg ttgtggaaaa tgctggacc cagacaaggc ttggctgttc taaggaaagt 360
caagagagtg gtggccagtt tgatgagagg attgtcctca aggaaacgcc gttcccatga 420
tgttctgact gtgcaattcc taattttggg catgctgggc atgacaatcg cagctacggt 480
tcgcaaggaa agagacggca gtacggatcat acgcgaggaa ggtaaggatg ccgctaccca 540
agtgagagtg gaaaatggta cctgcgtcat tctggccacc gacatgggct cttggtgtga 600
tgatagcctt tcttatgagt gcgtaacct agatcaaggt gaggaacctg ttgacgttga 660
ttgcttctgc cgaaaactgg atggggtgta tctcgaatat ggacggtgtg gtaacaaga 720
aggaagcaga accagacgct cagtgttat accctccac gctcaaggag agctgaccgg 780
acggggacat aaatggttgg agggcgactc actccgaaca catttgacc gcgtcgaggg 840
ctgggtctgg aaaaatcggc tgttggccct cgctatggtg acagtcgttt ggctcaegct 900
ggagtctgtg gttactcgcg tggcagtgct ggtggtgctc ctctgtcttg cccctgtcta 960
cgcgtccagg tgtactcatt tggaaaacag agatttgtc accggcaccc aggggacgac 1020

```

-continued

---

```

tcgggtaacc ctggtgcttg aactgggtgg ttgcgttact attaccgctg agggcaaac 1080
ctctatggat gtgtggcttg atgcaatcta tcaggagaat cccgcacaaa ccaggaata 1140
ttgccttcac gcaaaagtgt ccgatacaaa ggtecgcggt aggtgcccaa caatgggacc 1200
ggccaccctg gcggaggaac atcagggagg tacagtgtgc aaacgggacc agagtgatag 1260
aggctggggg aatcactgcg gcctgttcgg caaaggaagt attgtcgctt gcgtcaaggc 1320
agcctgtgag gccaaaaaga aggctactgg gcacgtctat gacgccaca agatcgttta 1380
tacagtgaaa gtggaaccac acacagggga ttactgtgcg gccaacgaga ctcatcccg 1440
tcgaaaaacg gccagcttca ccgtgtcatc cgaaaagacc atcctcacta tgggggagta 1500
tggcgacggt tctctgctct gccgggtggc tagcggagtc gacctggccc agacagtc 1560
cctggaactg gataaaacag ttgagcatct gcctaccgct tggcaggtgc acagggattg 1620
gtttaacgac cttgccctgc catggaaca tgaaggagcg agaaactgga ataatgcaga 1680
gcgactcgta gaattcgggt cccctcatgc cgtgaagatg gacgtctaca atctgggtga 1740
tcagaccggc gttctcctta aagctctcgc tggcgtacca gttgccaca tcgaaggaa 1800
gaagtaccac ctgaagtcag gccatgtaac ttgcgaggtg ggctggaga agttgaaat 1860
gaaaggtctt acgtacacaa tgtgtgacaa gaccaagttc acatggaaga gggcccccac 1920
agatagcggc cacgatactg tggatgga ggtgacctt tctggaacaa aaccctgcag 1980
aataccctg cgggctgtag ctcacggatc tcccgatgc aatgttgcta tgctgattac 2040
acctaaccct accatcgaga ataacgggtg tggttttatt gagatgcagc tccgccagc 2100
cgataacatc atctacgtgg gcgaactctc ttaccagtgg ttccagaaag ggagttcaat 2160
tgggctgggc ttccaaaaaa cgaagaaggg aatcgaacga ttgacgggta tcggcgagca 2220
cgcatgggat tttggttccg cagggggatt cctgtctct attggttaagg cactgcatac 2280
cgtgctgggg ggcgcattca attctatctt cggggcgctg gggttcctgc ctaaaactct 2340
gctgggagta gccctggcct ggttgggact gaatatcgcg aatccgacga tgtccatgct 2400
attctcttg gccggcgtgc ttgtactggc catgacactg ggcgttggcg ccgatcaagg 2460
atgcgccatc aactttggca agagagagct c 2491

```

&lt;210&gt; SEQ ID NO 51

&lt;211&gt; LENGTH: 2491

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 51

```

tcatttagga cacacgatga actccacgta accagacgtt tagctcaacg atccgttatt 60
tgtgtaaac taattaaaat tagcaagcaa ctcgctaate gtctcttgac tggctctgta 120
cagaccagca tttcgagtcc ctttttggga cccgcagtta taccatgctg ctctcaage 180
gaggaacagt ttgttttatt ttgtttttg ttttgttaa ctttgtctg gacctccaca 240
agttcctaaa tagaaaaaga aaaacaagtt gtaaaactga cttttttct agtgcgggt 300
ggatttctcc aacacctttt acgacctggg ttctgttccg aaccgacaag attcctttca 360
gttctctcac caccggcaaa actactctcc taacaggagt tcctttgctg caagggtact 420
acaagactga cacgttaagg attaaaacc gtacgaccog tactgtttagc gtegatgcca 480
agcgttctt tctctgctg catgccagta tgccgcctt ccattctac ggcgatgggt 540
tactctcac cttttaccat ggacgcagta agaccgggtg ctgtaccga gaaccacact 600

```

-continued

---

```

actatcggaa agaatactca cgcattggta tctagttcca ctcttggac aactgcaact 660
aacgaagacg gctttgcacc taccccacat agagcttata cctgccacac catttgttct 720
tccttcgtct tggctcgcga gtcacgaata tgggaggggtg cgagttcctc tcgactggcc 780
tgcccctgta tttaccaacc tcccgtgag tgaggcttgt gtaaactggg cgcagctccc 840
gaccagacc ttttagccg acaaccggga gcgataccac tgtcagcaaa ccgagtgcga 900
cctcagacac caatgagcgc accgtcacga ccaccacgag gagacagaac ggggacagat 960
gcgcaggctc acatgagtaa accttttgc tctaaaacag tggcctggg tcccctgctg 1020
agcccattgg gaccacgaac ttgaccacc aacgcaatga taatggcgac tcccgtttgg 1080
gagataccta cacaccgacc tacgtagat agtcctctta gggcgtgttt ggtcccttat 1140
aacggaagtg cgtttcgaca ggtatgttt ccagcgcga tccacgggtt gttaccctgg 1200
ccggtgggac cgcctccttg tagtccctc atgtcacacg tttgcctgg tctcactatc 1260
tccgacccca ttagtgcgc cggacaagcc gtttcttca taacagcgaa cgcagttccg 1320
tcggacactc cggttttct tccgatgacc cgtgcagata ctgcggttgt tctagcaaat 1380
atgtcacttt caccttgggtg tgtgtccct aatgcaccgc cggttgctct gagtaaggcc 1440
agcgttttgc cggtcgaagt ggcacagtag gcttttctgg taggagtgat accccctcat 1500
accgctgcaa agagacgaga cggcccaccg atcgcctcag ctggaccggg tctgtcagta 1560
ggaccttgac ctattttgc aactcgtaga cggatggcga accgtccacg tgtccctaac 1620
caaattgctg gaaccggacg gtaccttgt acttctcgc tctttgacct tattacgtct 1680
cgctgagcat cttaaagccac ggggagtacg gcaactctac ctgcagatgt tagaccact 1740
agtctggcgc caagaggaat ttcgagagcg accgcattgt caacgggtgt agcttccttg 1800
cttcatgggtg gacttcagtc cggtagattg aacgctccac ccggacctct tcaactttta 1860
ctttccagaa tgcattgttt acacactgtt ctggttcaag tgtacctct cccgggggtg 1920
tctatcgccg gtgctatgac accactacct ccaactgaaa agacctgtt ttgggacgtc 1980
ttatgggac gcccgacatc gagtgctag agggctacag ttacaacgat acgactaatg 2040
tggattggga tggtagctct tattgccacc accaaaataa ctctacgtcg aaggcggctc 2100
gctattgtag tagatgcacc cgcttgagag aatggtcacc aaagtcttc cctcaagtta 2160
accgcccag aaggttttt gcttcttccc ttagcttct aactgccaat agcgcctctg 2220
gcgtacccta aaaccaaggc gtcccctaa ggacagaaga taaccattcc gtgacgtatg 2280
gcacgacccc ccgcgtaagt taagataaaa gccccgcac cccaaggacg gatttgagga 2340
cgacctcat cgggaccgga ccaacctga cttatacgc ttaggctgct acaggtacag 2400
taaggagaac cggccgcacg aacatgaccg gtactgtgac ccgcaaccgc ggetagtctc 2460
tacgcggtag ttgaaaccgt tctctctcga g 2491

```

&lt;210&gt; SEQ ID NO 52

&lt;211&gt; LENGTH: 730

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 52

```

Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Tyr Leu
1 5 10 15

```

```

Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Arg

```

-continued

| 20  |     |     |     | 25  |     |     |     | 30  |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Ser | Lys | Gln | Lys | Lys | Arg | Gly | Gly | Thr | Asp | Trp | Met | Ser | Trp | Leu |
|     | 35  |     |     |     |     | 40  |     |     |     |     |     | 45  |     |     |     |
| Leu | Val | Ile | Gly | Met | Leu | Gly | Met | Thr | Ile | Ala | Ala | Thr | Val | Arg | Lys |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Glu | Arg | Asp | Gly | Ser | Thr | Val | Ile | Arg | Ala | Glu | Gly | Lys | Asp | Ala | Ala |
|     | 65  |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Thr | Gln | Val | Arg | Val | Glu | Asn | Gly | Thr | Cys | Val | Ile | Leu | Ala | Thr | Asp |
|     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |     |     | 95  |
| Met | Gly | Ser | Trp | Cys | Asp | Asp | Ser | Leu | Ser | Tyr | Glu | Cys | Val | Thr | Ile |
|     |     |     | 100 |     |     |     |     |     | 105 |     |     |     |     |     | 110 |
| Asp | Gln | Gly | Glu | Glu | Pro | Val | Asp | Val | Asp | Cys | Phe | Cys | Arg | Asn | Val |
|     |     |     | 115 |     |     |     |     |     | 120 |     |     |     |     |     | 125 |
| Asp | Gly | Val | Tyr | Leu | Glu | Tyr | Gly | Arg | Cys | Gly | Lys | Gln | Glu | Gly | Ser |
|     |     |     | 130 |     |     |     |     |     |     |     | 140 |     |     |     |     |
| Arg | Thr | Arg | Arg | Ser | Val | Leu | Ile | Pro | Ser | His | Ala | Gln | Gly | Glu | Leu |
|     |     |     |     |     | 150 |     |     |     |     |     | 155 |     |     |     | 160 |
| Thr | Gly | Arg | Gly | His | Lys | Trp | Leu | Glu | Gly | Asp | Ser | Leu | Arg | Thr | His |
|     |     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |
| Leu | Thr | Arg | Val | Glu | Gly | Trp | Val | Trp | Lys | Asn | Arg | Leu | Leu | Ala | Leu |
|     |     |     | 180 |     |     |     |     |     |     | 185 |     |     |     |     | 190 |
| Ala | Met | Val | Thr | Val | Val | Trp | Leu | Thr | Leu | Glu | Ser | Val | Val | Thr | Arg |
|     |     |     | 195 |     |     |     |     |     | 200 |     |     |     |     |     | 205 |
| Val | Ala | Val | Leu | Val | Val | Leu | Leu | Cys | Leu | Ala | Pro | Val | Tyr | Ala | Ser |
|     |     |     | 210 |     |     |     |     |     | 215 |     |     |     |     |     | 220 |
| Arg | Cys | Thr | His | Leu | Glu | Asn | Arg | Asp | Phe | Val | Thr | Gly | Thr | Gln | Gly |
|     |     |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     | 240 |
| Thr | Thr | Arg | Val | Thr | Leu | Val | Leu | Glu | Leu | Gly | Gly | Cys | Val | Thr | Ile |
|     |     |     |     |     |     | 245 |     |     |     |     | 250 |     |     |     | 255 |
| Thr | Ala | Glu | Gly | Lys | Pro | Ser | Met | Asp | Val | Trp | Leu | Asp | Ala | Ile | Tyr |
|     |     |     |     |     |     |     |     |     | 265 |     |     |     |     |     | 270 |
| Gln | Glu | Asn | Pro | Ala | Gln | Thr | Arg | Glu | Tyr | Cys | Leu | His | Ala | Lys | Leu |
|     |     |     |     |     |     |     |     |     |     |     | 280 |     |     |     | 285 |
| Ser | Asp | Thr | Lys | Val | Ala | Ala | Arg | Cys | Pro | Thr | Met | Gly | Pro | Ala | Thr |
|     |     |     |     |     |     |     |     |     |     |     | 290 |     |     |     | 300 |
| Leu | Ala | Glu | Glu | His | Gln | Gly | Gly | Thr | Val | Cys | Lys | Arg | Asp | Gln | Ser |
|     |     |     |     |     |     |     |     |     |     |     | 310 |     |     |     | 320 |
| Asp | Arg | Gly | Trp | Gly | Asn | His | Cys | Gly | Leu | Phe | Gly | Lys | Gly | Ser | Ile |
|     |     |     |     |     |     |     |     |     |     |     | 325 |     |     |     | 335 |
| Val | Ala | Cys | Val | Lys | Ala | Ala | Cys | Glu | Ala | Lys | Lys | Lys | Ala | Thr | Gly |
|     |     |     |     |     |     |     |     |     |     |     | 340 |     |     |     | 350 |
| His | Val | Tyr | Asp | Ala | Asn | Lys | Ile | Val | Tyr | Thr | Val | Lys | Val | Glu | Pro |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 355 |
| His | Thr | Gly | Asp | Tyr | Val | Ala | Ala | Asn | Glu | Thr | His | Ser | Gly | Arg | Lys |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 370 |
| Thr | Ala | Ser | Phe | Thr | Val | Ser | Ser | Glu | Lys | Thr | Ile | Leu | Thr | Met | Gly |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 385 |
| Glu | Tyr | Gly | Asp | Val | Ser | Leu | Leu | Cys | Arg | Val | Ala | Ser | Gly | Val | Asp |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 405 |
| Leu | Ala | Gln | Thr | Val | Ile | Leu | Glu | Leu | Asp | Lys | Thr | Val | Glu | His | Leu |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 420 |
| Pro | Thr | Ala | Trp | Gln | Val | His | Arg | Asp | Trp | Phe | Asn | Asp | Leu | Ala | Leu |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 435 |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 440 |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 445 |

-continued

Pro Trp Lys His Glu Gly Ala Arg Asn Trp Asn Asn Ala Glu Arg Leu  
 450 455 460  
 Val Glu Phe Gly Ala Pro His Ala Val Lys Met Asp Val Tyr Asn Leu  
 465 470 475 480  
 Gly Asp Gln Thr Gly Val Leu Leu Lys Ala Leu Ala Gly Val Pro Val  
 485 490 495  
 Ala His Ile Glu Gly Thr Lys Tyr His Leu Lys Ser Gly His Val Thr  
 500 505 510  
 Cys Glu Val Gly Leu Glu Lys Leu Lys Met Lys Gly Leu Thr Tyr Thr  
 515 520 525  
 Met Cys Asp Lys Thr Lys Phe Thr Trp Lys Arg Ala Pro Thr Asp Ser  
 530 535 540  
 Gly His Asp Thr Val Val Met Glu Val Thr Phe Ser Gly Thr Lys Pro  
 545 550 555 560  
 Cys Arg Ile Pro Val Arg Ala Val Ala His Gly Ser Pro Asp Val Asn  
 565 570 575  
 Val Ala Met Leu Ile Thr Pro Asn Pro Thr Ile Glu Asn Asn Gly Gly  
 580 585 590  
 Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr Val  
 595 600 605  
 Gly Glu Leu Ser Tyr Gln Trp Phe Gln Lys Gly Ser Ser Ile Gly Arg  
 610 615 620  
 Val Phe Gln Lys Thr Lys Lys Gly Ile Glu Arg Leu Thr Val Ile Gly  
 625 630 635 640  
 Glu His Ala Trp Asp Phe Gly Ser Ala Gly Gly Phe Leu Ser Ser Ile  
 645 650 655  
 Gly Lys Ala Leu His Thr Val Leu Gly Gly Ala Phe Asn Ser Ile Phe  
 660 665 670  
 Gly Gly Val Gly Phe Leu Pro Lys Leu Leu Leu Gly Val Ala Leu Ala  
 675 680 685  
 Trp Leu Gly Leu Asn Met Arg Asn Pro Thr Met Ser Met Ser Phe Leu  
 690 695 700  
 Leu Ala Gly Val Leu Val Leu Ala Met Thr Leu Gly Val Gly Ala Asp  
 705 710 715 720  
 Thr Gly Cys Ala Ile Asp Ile Ser Arg Gln  
 725 730

&lt;210&gt; SEQ ID NO 53

&lt;211&gt; LENGTH: 2286

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 53

```

agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta 60
acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc 120
ggcaagagcc gggctgtcta tttgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt 180
ggacttaagc ggagctccaa acaaaagaaa cgggggggaa cagactggat gagctggctg 240
ctcgtaatcg gcatgctggg catgacaatc gcagctacgg ttcgcaagga aagagacggc 300
agtacggtca tacgcgcca aggtaaggat gccgctaccc aagtgagagt ggaaaatggt 360
acctgcgtca ttctggccac cgacatgggc tcttggtgtg atgatagcct ttcttatgag 420

```

-continued

---

|                                                                      |      |
|----------------------------------------------------------------------|------|
| tgcgtaacca tagatcaagg tgaggaacct gttgacgttg attgcttctg ccgaaacgtg    | 480  |
| gatgggggtgt atctcgaata tggacgggtgt ggtaaacaaag aaggaagcag aaccagacgc | 540  |
| tcagtgtcta taccctccca cgctcaagga gagctgaccg gacggggaca taaatggttg    | 600  |
| gagggcgact cactccgaac acatttgacc cgcgtcgagg gctgggtctg gaaaaatcgg    | 660  |
| ctggtggccc tcgctatggt gacagtcggt tggctcacgc tggagtctgt ggttactcgc    | 720  |
| gtggcagtcg tgggtgtgct cctctgtctt gccctgtct acgcgtccag gtgtactcat     | 780  |
| ttggaaaaa gagatthtgt caccggcacc caggggacga ctccggtaac cctggtgctt     | 840  |
| gaactgggtg gttcgcgttac tattaccgct gagggcaaac cctctatgga tgtgtggctg   | 900  |
| gatgcaatct atcaggagaa tcccgcacaa accagggaat attgccttca cgcaaacgtg    | 960  |
| tccgatacaa aggtcgcggc taggtgocca acaatgggac cggccacctt ggcggaggaa    | 1020 |
| catcagggag gtacagtgtg caaacgggac cagagtgata gaggctgggg taactactgc    | 1080 |
| ggcctgttcg gcaaaaggaag tattgtcgtc tgcgtcaagg cagcctgtga ggccaaaaag   | 1140 |
| aaggctactg ggcacgtcta tgacgccaac aagatcgttt atacagtgaa agtggaaacca   | 1200 |
| cacacagggg attacgtggc ggccaacgag actcattccg gtcgcaaac ggccagcttc     | 1260 |
| accgtgtcat ccgaaaagac catcctcact atgggggagt atggcgacgt ttctctgctc    | 1320 |
| tgccgggtgg ctagcggagt cgacctggcc cagacagtca tcctggaact ggataaaaca    | 1380 |
| gttgagcctc tgcctaccgc ttggcagggt cacagggatt ggtttaacga ccttgccctg    | 1440 |
| ccatggaaac atgaaggagc gagaaaactgg aataatgcag agcgactcgt agaattcggg   | 1500 |
| gcccctcatg ccgtgaagat ggacgtctac aatctgggtg atcagaccgg cgttctcctt    | 1560 |
| aaagctctcg ctggcgtacc agttgcccac atcgaaggaa cgaagtacca cctgaagtca    | 1620 |
| ggccatgtaa cttcggaggt gggcctggag aagttgaaaa tgaaaggctt tacgtacaca    | 1680 |
| atgtgtgaca agaccaagtt cacatggaag agggccccc cagatagcgg ccacgatact     | 1740 |
| gtggtgatgg aggtgacctt ttctggaaca aaacctgca gaatacccg gcggtgtgta      | 1800 |
| gctcacggat ctcccgatgt caatgttctc atgctgatta cacctaacc taccatcgag     | 1860 |
| aataacggtg gtggttttat tgagatgcag cttccgccag gcgataacat catctactg     | 1920 |
| ggcgaactct cttaccagtg gtttcagaaa gggagttcaa ttgggcgggt cttccaaaaa    | 1980 |
| acgaagaagg gaatcgaacg attgacgggt atcggcgagc acgcatggga ttttggttcc    | 2040 |
| gcaggggat tcctgtcttc tatttgtaag gcaactgcata ccgtgctggg gggcgcattc    | 2100 |
| aattctatth tcggggcgct ggggttctc cctaaactcc tgctgggagt agccctggcc     | 2160 |
| tggttgggac tgaatatgag gaatccgacg atgtccatgt cattcctctt ggcggcgctg    | 2220 |
| cttgacttgg ccatgacact gggcgttggc gccgacactg ggtgtgcat agacatcagc     | 2280 |
| cggcaa                                                               | 2286 |

&lt;210&gt; SEQ ID NO 54

&lt;211&gt; LENGTH: 2286

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 54

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| tcateaacgc gacacactcg actgthttaa tcaacacaaa cactcctaatt tgtgtttaat | 60  |
| tgtgtcacgc tcgacaaaaga atcgtgcttc tagagctaca gattctttgg tcctcccggg | 120 |
| ccgttctcgg cccgacagat aaacgatttt gcgccttacg gggcgacaaa caggaactaa  | 180 |

-continued

---

```

cctgaattcg cctegaggtt tgttttcttt gccccccctt gtctgacctt ctegaccgac 240
gagcattagc cgtacgacct gtactgttag cgtegatgcc aagcgttccct ttctctgccg 300
tcatgccagt atgcgcgcct tccattccta cggcgatggg ttcactctca ccttttacca 360
tggacgcagt aagaccggtg gctgtaccgg agaaccacac tactatcgga aagaatactc 420
acgcattggg atctagtacc actccttggg caactgcaac taacgaagac ggctttgcac 480
ctaccccaca tagagcttat acctgccaca ccatttgctt ttccttcgtc ttggtctgcg 540
agtcacgaat atgggagggg gcgagttcct ctgcactggc ctgcccctgt atttaccac 600
ctcccgtga gtgaggcttg tgtaaactgg gcgcagctcc cgaccagac ctttttagcc 660
gacaaccggg agcgatacca ctgtcagcaa accgagtgcg acctcagaca ccaatgagcg 720
cacgctcacg accaccacga ggagacagaa cggggacaga tgcgcaggtc cacatgagta 780
aaccttttgt ctctaaaaca gtggccgtgg gtcccctgct gagcccattg ggaccacgaa 840
cttgaccac caacgcaatg ataatggcga ctcccgtttg ggagatacct acacaccgac 900
ctacgttaga tagtcctctt agggcgtgtt tgggtccctt taacggaagt gcgtttcgac 960
aggctatggt tccagcgccg atccacgggt tgttacctg gccggtggga ccgctcctt 1020
gtagtccctc catgtcacac gtttgccctg gtctcactat ctccgacccc attagtgacg 1080
ccggacaagc cgtttccttc ataacagcga acgcagttcc gteggacact cgggttttcc 1140
ttccgatgac ccgtgcagat actgcggttg ttctagcaaa tatgtcactt tcacctggt 1200
gtgtgtcccc taatgcaccg ccggttgcctc tgagtaaggc cagcgttttg ccggtcgaag 1260
tggcacagta ggctttctct gtaggagtga taccocctca taccgctgca aagagacgag 1320
acggcccacc gatcgctca gctggaccgg gtctgtcagt aggacctga cctatttgt 1380
caactcgtag acggatggcg aaccgtccac gtgtccctaa ccaaattgct ggaacgggac 1440
ggtaccttgg tacttctctg ctctttgacc ttattacgtc tcgctgagca tcttaagcca 1500
cggggagtac ggcacttcta cctgcagatg ttagaccac tagtctggcc gcaagaggaa 1560
tttcgagagc gaccgcatgg tcaacgggtg tagcttccct gcttcattgg ggacttcagt 1620
ccggtacatt gaacgctcca ccgggaacct ttcaactttt actttccaga atgcattggt 1680
tacacactgt tctggttcaa gtgtacctc tcccgggggt gtctatcgcc ggtgctatga 1740
caccactacc tccactggaa aagacctgtt ttgggaagct cttatgggca cgcccacat 1800
cgagtgccta gagggctaca gttacaacga tacgactaat gtggattggg atggtagctc 1860
ttattgccac caccaaaata actctacgtc gaaggcggtc cgctattgta gtagatgcac 1920
ccgcttgaga gaattgctac caaagtcttt cctcaagtt aaccgcccc gaaggttttt 1980
tgcttcttcc cttagcttgc taactgccc tagccgctcg tgcgtacctt aaaaccaagg 2040
cgtcccccta aggacagaag ataaccattc cgtgacgtat ggcacgacct cccgcgtaag 2100
ttaagataaa agcccccgca cccaaggac ggatttgagg acgacctca tcgggaccgg 2160
accaacctg acttatacgc cttaggctgc tacaggta gtaaggagaa cgggcccgcac 2220
gaacatgacc ggtactgtga cccgcaaccg cggctgtgac ccacacggta tctgtagtgc 2280
gccgtt 2286

```

&lt;210&gt; SEQ ID NO 55

&lt;211&gt; LENGTH: 727

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

-continued

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 55

```

Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Tyr Leu
 1 5 10 15
Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Arg
 20 25 30
Ser Ser Lys Gln Lys Lys Arg Gly Gly Lys Thr Gly Ile Ala Val Met
 35 40 45
Ile Gly Met Leu Ala Cys Val Gly Ala Ala Thr Val Arg Lys Glu Arg
 50 55 60
Asp Gly Ser Thr Val Ile Arg Ala Glu Gly Lys Asp Ala Ala Thr Gln
 65 70 75 80
Val Arg Val Glu Asn Gly Thr Cys Val Ile Leu Ala Thr Asp Met Gly
 85 90 95
Ser Trp Cys Asp Asp Ser Leu Ser Tyr Glu Cys Val Thr Ile Asp Gln
 100 105 110
Gly Glu Glu Pro Val Asp Val Asp Cys Phe Cys Arg Asn Val Asp Gly
 115 120 125
Val Tyr Leu Glu Tyr Gly Arg Cys Gly Lys Gln Glu Gly Ser Arg Thr
 130 135 140
Arg Arg Ser Val Leu Ile Pro Ser His Ala Gln Gly Glu Leu Thr Gly
 145 150 155 160
Arg Gly His Lys Trp Leu Glu Gly Asp Ser Leu Arg Thr His Leu Thr
 165 170 175
Arg Val Glu Gly Trp Val Trp Lys Asn Arg Leu Leu Ala Leu Ala Met
 180 185 190
Val Thr Val Val Trp Leu Thr Leu Glu Ser Val Val Thr Arg Val Ala
 195 200 205
Val Leu Val Val Leu Leu Cys Leu Ala Pro Val Tyr Ala Ser Arg Cys
 210 215 220
Thr His Leu Glu Asn Arg Asp Phe Val Thr Gly Thr Gln Gly Thr Thr
 225 230 235 240
Arg Val Thr Leu Val Leu Glu Leu Gly Gly Cys Val Thr Ile Thr Ala
 245 250 255
Glu Gly Lys Pro Ser Met Asp Val Trp Leu Asp Ala Ile Tyr Gln Glu
 260 265 270
Asn Pro Ala Gln Thr Arg Glu Tyr Cys Leu His Ala Lys Leu Ser Asp
 275 280 285
Thr Lys Val Ala Ala Arg Cys Pro Thr Met Gly Pro Ala Thr Leu Ala
 290 295 300
Glu Glu His Gln Gly Gly Thr Val Cys Lys Arg Asp Gln Ser Asp Arg
 305 310 315 320
Gly Trp Gly Asn His Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala
 325 330 335
Cys Val Lys Ala Ala Cys Glu Ala Lys Lys Lys Ala Thr Gly His Val
 340 345 350
Tyr Asp Ala Asn Lys Ile Val Tyr Thr Val Lys Val Glu Pro His Thr
 355 360 365
Gly Asp Tyr Val Ala Ala Asn Glu Thr His Ser Gly Arg Lys Thr Ala
 370 375 380
Ser Phe Thr Val Ser Ser Glu Lys Thr Ile Leu Thr Met Gly Glu Tyr
 385 390 395 400

```

-continued

---

Gly Asp Val Ser Leu Leu Cys Arg Val Ala Ser Gly Val Asp Leu Ala  
 405 410 415

Gln Thr Val Ile Leu Glu Leu Asp Lys Thr Val Glu His Leu Pro Thr  
 420 425 430

Ala Trp Gln Val His Arg Asp Trp Phe Asn Asp Leu Ala Leu Pro Trp  
 435 440 445

Lys His Glu Gly Ala Arg Asn Trp Asn Asn Ala Glu Arg Leu Val Glu  
 450 455 460

Phe Gly Ala Pro Ala Val Lys Met Asp Val Tyr Asn Leu Gly Asp Gln  
 465 470 475 480

Thr Gly Val Leu Leu Lys Ala Leu Ala Gly Val Pro Val Ala His Ile  
 485 490 495

Glu Gly Thr Lys Tyr His Leu Lys Ser Gly His Val Thr Cys Glu Val  
 500 505 510

Gly Leu Glu Lys Leu Lys Met Lys Gly Leu Thr Tyr Thr Met Cys Asp  
 515 520 525

Lys Thr Lys Phe Thr Trp Lys Arg Ala Pro Thr Asp Ser Gly His Asp  
 530 535 540

Thr Val Val Met Glu Val Thr Phe Ser Gly Thr Lys Pro Cys Arg Ile  
 545 550 555 560

Pro Val Arg Ala Val Ala His Gly Ser Pro Asp Val Asn Val Ala Met  
 565 570 575

Leu Ile Thr Pro Asn Pro Thr Ile Glu Asn Asn Gly Gly Gly Phe Ile  
 580 585 590

Glu Met Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr Val Gly Glu Leu  
 595 600 605

Ser Tyr Gln Trp Phe Gln Lys Gly Ser Ser Ile Gly Arg Val Phe Gln  
 610 615 620

Lys Thr Lys Lys Gly Ile Glu Arg Leu Thr Val Ile Gly Glu His Ala  
 625 630 635 640

Trp Asp Phe Gly Ser Ala Gly Gly Phe Leu Ser Ser Ile Gly Lys Ala  
 645 650 655

Leu His Thr Val Leu Gly Gly Ala Phe Asn Ser Ile Phe Gly Gly Val  
 660 665 670

Gly Phe Leu Pro Lys Leu Leu Leu Gly Val Ala Leu Ala Trp Leu Gly  
 675 680 685

Leu Asn Met Arg Asn Pro Thr Met Ser Met Ser Phe Leu Leu Ala Gly  
 690 695 700

Val Leu Val Leu Ala Met Thr Leu Gly Val Gly Ala Asp Thr Gly Cys  
 705 710 715 720

Ala Ile Asp Ile Ser Arg Gln  
 725

<210> SEQ ID NO 56  
 <211> LENGTH: 2280  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 56

agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta 60  
 acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc 120  
 ggcaagagcc gggctgtcta ttgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt 180

-continued

---

```

ggacttaagc ggagctccaa gcaaaagaaa cgcgggggaa agacaggcat agctgtgatg 240
ataggcatgc tggcttgtgt cggagcagct accgtgcgaa aagaacgcga cggaaagcacc 300
gtgataaggg ctgagggtaa ggatgaggct acgcagggtga gagtagagaa tggcacttgc 360
gtaatactcg cgactgatat gggatcctgg tgtgacgata gcctcagtta tgaatgcgta 420
acaatagacc agggcgaaga acctgtggac gttgactgtt tctgtagaaa tgtggatggc 480
gtttatctgg agtacggcgc ctgtggaaaa caggagggtc cacgaactcg aagatctgtg 540
ctgattccaa gtcacgcgca aggagagttg accggtagag gccacaagtg gcttgaaggg 600
gactcattga ggaccacact gactaggggtg gagggttggg tttggaagaa tcggttgctc 660
gogctcgcta tggtcaccgt cgtgtggctg acactggaga gtgtcgtgac tcgggttget 720
gtgttggttg tcctcctctg tttggcccca gtgtacgcgt ccagggtgac tcatttgaa 780
aacagagatt ttgtcacccg caccagggg acgactcggg taaccctggt gcttgaactg 840
ggtggttgcg ttactattac cgtgagggc aaacctcta tggatgtgtg gctggatgca 900
atctatcagg agaatccgcg acaaacagg gaatattgcc ttcacgcaa gctgtccgat 960
acaaaggtcg cggtaggtg cccaacaatg ggaccggcca ccctggcga ggaacatcag 1020
ggaggtacag tgtgcaaacg ggaccagagt gatagaggct ggggtaatca ctgcgccctg 1080
ttcggcaaa gaagtattgt cgcttgcgtc aaggcagcct gtgaggccaa aaagaaggct 1140
actgggcacg tctatgacgc caacaagatc gtttatacag tgaaagtgga accacacaca 1200
ggggattacg tggcgcccaa cgagactcat tccggtcgca aaacggccag cttcacctg 1260
tcatccgaaa agaccatcct cactatgggg gagtatggcg acgtttctct gctctgccg 1320
gtggctagcg gagtcgacct ggcccagaca gtcacctcgg aactggataa aacagttgag 1380
catctgccta ccgcttgga ggtgcacagg gattggttta acgacctgc cctgccatgg 1440
aaacatgaag gagcgagaaa ctggaataat gcagagcgac tcgtagaatt cggtgcccct 1500
catgccgtga agatggacgt ctacaatctg ggtgatcaga ccggcgttct ccttaaagct 1560
ctcgtggcg taccagttgc ccacatcgaa ggaacgaagt accacctgaa gtcaggccat 1620
gtaacttgcg aggtgggctt ggagaagttg aaaatgaaag gtcttacgta cacaatgtgt 1680
gacaagacca agttcacatg gaagagggcc cccacagata gcggccacga tactgtggtg 1740
atggagggtg ccttttctg aacaaaacc tgcagaatac ccgtgcgggc tgtagctcac 1800
ggatctcccg atgtcaatgt tgctatgctg attacaccta accctacat cgagaataac 1860
ggtgtggtt ttattgagat gcagcttccg ccaggcgata acatcatcta cgtgggcgaa 1920
ctctcttacc agtgggttca gaaagggagt tcaattgggc gggcttcca aaaaacgaag 1980
aaggaatcg aacgattgac ggttatcggc gagcacgcat gggattttgg ttccgcagg 2040
ggattcctgt cttctattgg taaggcactg cataccgtgc tggggggcgc attcaattct 2100
atcttcgggg gcgtggggtt cctgcctaaa ctctctgctg gagtagccct ggctgggtg 2160
ggactgaata tgcggaatcc gacgatgtcc atgtcattcc tcttggccgg cgtgcttcta 2220
ctggccatga cactgggctg tggcgcgac actgggtgtg ccatagacat cagccggcaa 2280

```

&lt;210&gt; SEQ ID NO 57

&lt;211&gt; LENGTH: 2280

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 57

-continued

---

|             |             |            |            |             |            |      |
|-------------|-------------|------------|------------|-------------|------------|------|
| tcatcaagcg  | gacacactcg  | actgtttgaa | tcatcacaaa | cactoctaat  | tgttgtaaat | 60   |
| tgtgtcacgc  | tcgacaaaga  | atcgtgcttc | tagagetaca | gattctttgg  | tcctcccggg | 120  |
| ccgttctcgg  | cccgacagat  | aaacgatttt | gcgccttacg | gggcgcaaaa  | caggaactaa | 180  |
| cctgaattcg  | cctcgagggt  | cgttttcttt | gcgccccctt | tctgtccgta  | tcgacactac | 240  |
| tatccgtaag  | accgaacaca  | gcctcgtcga | tggcacgctt | ttcttgcgct  | gccttcgtgg | 300  |
| cactattccc  | gactcccatt  | cctacgccga | tgcgtccact | ctcatctctt  | accgtgaacg | 360  |
| cattatgagc  | gctgactata  | ccctaggacc | acactgctat | cggagtcaat  | acttacgcat | 420  |
| tgttatctgg  | tcccgttctt  | tggacaactg | caactgacaa | agacatcttt  | acacctaccg | 480  |
| caaatagacc  | tcattgccggc | gacacctttt | gtcctcccga | gtgcttgagc  | ttctagacac | 540  |
| gactaaggtt  | cagtgcgcgt  | tcctctcaac | tggccatctc | cgggtgtcac  | cgaacttccc | 600  |
| ctgagtaact  | cctgggtgga  | ctgatcccac | ctcccacccc | aaaccttctt  | agccaacgag | 660  |
| cgcgagcgat  | accagtggca  | gcacaccgac | tgtgacctct | cacagcactg  | agcccaacga | 720  |
| cacaaccaac  | aggaggagac  | aaaccggggg | cacatgcgca | ggtccacatg  | agtaaacctt | 780  |
| ttgtctctaa  | aacagtggcc  | gtgggtcccc | tgctgagccc | attgggacca  | cgaacttgac | 840  |
| ccaccaacgc  | aatgataaat  | gcgactcccc | tttgggagat | acctacacac  | cgacctacgt | 900  |
| tagatagtcc  | tcttagggcg  | tgtttggtcc | cttataacgg | aagtgcgttt  | cgacaggcta | 960  |
| tgtttccagc  | gccgatccac  | gggttggtac | cctggccggg | gggaccgcct  | ccttgtagt  | 1020 |
| cctccatgtc  | acacgtttgc  | cctggtctca | ctatctccga | ccccattagt  | gacgcccggc | 1080 |
| aagccgtttc  | cttcataaca  | gcgaacgcag | ttccgtcgga | cactccgggt  | tttcttccga | 1140 |
| tgaccctgtc  | agatactgcg  | gttgttctag | caaatatgtc | actttcacct  | tggtgtgtgt | 1200 |
| cccctaattg  | accgccgggt  | gctctgagta | aggccagcgt | tttgccggtc  | gaagtggcac | 1260 |
| agtaggcttt  | tctggttaga  | gtgatacccc | ctcataccgc | tgcaaaagaga | cgagacggcc | 1320 |
| caccgatcgc  | ctcagctgga  | ccgggtctgt | cagtaggacc | ttgacctatt  | ttgtcaactc | 1380 |
| gtagacggat  | ggcgaaccgt  | ccacgtgtcc | ctaaccacaa | tgctggaacg  | ggacggtaac | 1440 |
| tttgtacttc  | ctcgtctctt  | gaccttatta | cgtctcgtcg | agcatcttaa  | gccacgggga | 1500 |
| gtacggcact  | tctacctgca  | gatgtagac  | ccactagtct | ggccgcaaga  | ggaatttcga | 1560 |
| gagcgaccgc  | atggtcaacg  | ggtgtagctt | ccttgcctca | tggtggactt  | cagtccggta | 1620 |
| cattgaacgc  | tcaccccgga  | cctcttcaac | ttttactttc | cagaatgcat  | gtgttacaca | 1680 |
| ctgttctggg  | tcaagtgtac  | cttctcccgg | gggtgtctat | cgccgggtgt  | atgacaccac | 1740 |
| tacctccact  | ggaaaagacc  | ttgttttggg | acgtcttatg | ggcacgcccc  | acatcgagt  | 1800 |
| cctagagggc  | tacagttaca  | acgatacgac | taatgtggat | tgggatggta  | gctcttattg | 1860 |
| ccaccaccaa  | aataactcta  | cgtcgaaggc | ggtccgctat | tgtagtagat  | gcacccgctt | 1920 |
| gagagaatgg  | tcaccaaagt  | ctttccctca | agttaaccgg | cccagaaggt  | tttttgcttc | 1980 |
| ttcccttagc  | ttgctaactg  | ccaatagccg | ctcgtgcgta | ccctaaaacc  | aaggcgtccc | 2040 |
| cctaaggaca  | gaagataacc  | attccgtgac | gtatggcacg | accccccgcg  | taagttaaga | 2100 |
| taaaagcccc  | cgccccccaa  | ggacggattt | gaggacgacc | ctcatcgggg  | ccggaccaac | 2160 |
| cctgacttat  | acgccttagg  | ctgctacagg | tacagtaagg | agaaccggcc  | gcacgaacat | 2220 |
| gaccggctact | gtgaccgcga  | accgcggctg | tgacccacac | ggtatctgta  | gtcggccggt | 2280 |

-continued

---

```

<211> LENGTH: 635
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 58
Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Tyr Leu
1 5 10 15
Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Gln
20 25 30
Lys Lys Arg Gly Gly Lys Thr Gly Ile Ala Val Ile Val Pro Gln Ala
35 40 45
Leu Leu Phe Val Pro Leu Leu Val Phe Pro Leu Cys Phe Gly Lys Phe
50 55 60
Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro Trp Ser Pro Ile Asp
65 70 75 80
Ile His His Leu Ser Cys Pro Asn Asn Leu Val Val Glu Asp Glu Gly
85 90 95
Cys Thr Asn Leu Ser Gly Phe Ser Tyr Met Glu Leu Lys Val Gly Tyr
100 105 110
Ile Ser Ala Ile Lys Met Asn Gly Phe Thr Cys Thr Gly Val Val Thr
115 120 125
Glu Ala Glu Thr Tyr Thr Asn Phe Val Gly Tyr Val Thr Thr Thr Phe
130 135 140
Lys Arg Lys His Phe Arg Pro Thr Pro Asp Ala Cys Arg Ala Ala Tyr
145 150 155 160
Asn Trp Lys Met Ala Gly Asp Pro Arg Tyr Glu Glu Ser Leu His Asn
165 170 175
Pro Tyr Pro Asp Tyr His Trp Leu Arg Thr Val Lys Thr Thr Lys Glu
180 185 190
Ser Leu Val Ile Ile Ser Pro Ser Val Ala Asp Leu Asp Pro Tyr Asp
195 200 205
Arg Ser Leu His Ser Arg Val Phe Pro Gly Gly Asn Cys Ser Gly Val
210 215 220
Ala Val Ser Ser Thr Tyr Cys Ser Thr Asn His Asp Tyr Thr Ile Trp
225 230 235 240
Met Pro Glu Asn Pro Arg Leu Gly Met Ser Cys Asp Ile Phe Thr Asn
245 250 255
Ser Arg Gly Lys Arg Ala Ser Lys Gly Ser Glu Thr Cys Gly Phe Val
260 265 270
Asp Glu Arg Gly Leu Tyr Lys Ser Leu Lys Gly Ala Cys Lys Leu Lys
275 280 285
Leu Cys Gly Val Leu Gly Leu Arg Leu Met Asp Gly Thr Trp Val Ala
290 295 300
Met Gln Thr Ser Asn Glu Thr Lys Trp Cys Pro Pro Gly Gln Leu Val
305 310 315 320
Asn Leu His Asp Phe Arg Ser Asp Glu Ile Glu His Leu Val Val Glu
325 330 335
Glu Leu Val Lys Lys Arg Glu Glu Cys Leu Asp Ala Leu Glu Ser Ile
340 345 350
Met Thr Thr Lys Ser Val Ser Phe Arg Arg Leu Ser His Leu Arg Lys
355 360 365
Leu Val Pro Gly Phe Gly Lys Ala Tyr Thr Ile Phe Asn Lys Thr Leu
370 375 380

```

-continued

Met Glu Ala Asp Ala His Tyr Lys Ser Val Arg Thr Trp Asn Glu Ile  
 385 390 395 400

Ile Pro Ser Lys Gly Cys Leu Arg Val Gly Gly Arg Cys His Pro His  
 405 410 415

Val Asn Gly Val Phe Phe Asn Gly Ile Ile Leu Gly Pro Asp Gly Asn  
 420 425 430

Val Leu Ile Pro Glu Met Gln Ser Ser Leu Leu Gln Gln His Met Glu  
 435 440 445

Leu Leu Val Ser Ser Val Ile Pro Leu Met His Pro Leu Ala Asp Pro  
 450 455 460

Ser Thr Val Phe Lys Asn Gly Asp Glu Ala Glu Asp Phe Val Glu Val  
 465 470 475 480

His Leu Pro Asp Val His Glu Arg Ile Ser Gly Val Asp Leu Gly Leu  
 485 490 495

Pro Asn Trp Gly Lys Tyr Val Leu Leu Ser Ala Gly Ala Leu Thr Ala  
 500 505 510

Leu Met Leu Ile Ile Phe Leu Met Thr Cys Trp Arg Arg Val Asn Arg  
 515 520 525

Ser Glu Pro Thr Gln His Asn Leu Arg Gly Thr Gly Arg Glu Val Ser  
 530 535 540

Val Thr Pro Gln Ser Gly Lys Ile Ile Ser Ser Trp Glu Ser Tyr Lys  
 545 550 555 560

Ser Gly Gly Glu Thr Gly Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly  
 565 570 575

Asp Val Glu Ser Asn Pro Gly Pro Ala Arg Asp Arg Ser Ile Ala Leu  
 580 585 590

Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser Val Asn Val  
 595 600 605

His Ala Asp Thr Gly Cys Ala Ile Asp Ile Ser Arg Gln Glu Leu Arg  
 610 615 620

Cys Gly Ser Gly Val Phe Ile His Asn Asp Val  
 625 630 635

&lt;210&gt; SEQ ID NO 59

&lt;211&gt; LENGTH: 2000

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 59

```

agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta 60
acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc 120
ggcaagagcc gggctgtcta tttgctaaaa cgcggaatgc cccgcgtggt gtccttgatt 180
ggacttaagc aaaagaagcg agggggcaag actggtatag ctgtgatcgt tcctcaggct 240
cttttgtttg tacccttgcg ggtatttccc ctttgctttg gtaaatttcc tatctatacc 300
atccctgata agctcggggc ttggagtecc attgatattc accatttgag ctgccc aaac 360
aacctcgtcg ttgaggatga aggggtgcact aatctttctg gattttccta catggagttg 420
aaagtgggct atatttcagc cattaagatg aacggcttta cttgtacagg agtcgtgacc 480
gaagccgaga catatacaaa tttcgtggga tacgtcacca ccaccttcaa gagaaaaaac 540
ttccgccc aa cgctgacgc ttgtcggggc gcttacaact ggaagatggc aggagatcct 600

```

-continued

---

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cgatatgaag aatctctgca caaccgcat cctgattacc attggctgcg gacagtcaag   | 660  |
| actaccaagg agagtctggt cattatatca ccaagcgtgg cogatcttga tccttatgat  | 720  |
| agatccctgc acagtagggt ttttctggc ggggaattgta gcggtgttgc agtatcaagt  | 780  |
| acctactgct ccactaacca cgactacact atatggatgc ctgagaacct tcgactcggg  | 840  |
| atgagttgcg acatttttac gaactcacgg ggcaagcggg catctaagggt gtctgaaca  | 900  |
| tgcggtttg ttgatgagcg ggggtgtgat aaatctctta aaggcgcctg taagctgaaa   | 960  |
| ctctgtggcg tactggggct gcgcctgatg gacggcacat ggggtgctat gcagacaagc  | 1020 |
| aatgaaacaa agtgggtgcc cctcggtcag ctggttaatc tgcacgactt taggtctgac  | 1080 |
| gaaatcgagc accttgggtt ggaggaactg gtgaagaaac gcgaagagtg cctggacgca  | 1140 |
| cttgagagta ttatgaccac caaatccgtt tccttcagaa gactgagcca cctgcgaaaag | 1200 |
| ctggtgccag ggttcgggaa ggcttatact atttcaaca agactcttat ggaggcggat   | 1260 |
| gcccattata agtcagttag gacttgggat gagataatc cctccaaagg atgtctgaga   | 1320 |
| gtcggtgagg gatgccacc ccatgtcaat ggggtgttct ttaacggaat catcctggga   | 1380 |
| cctgacggga acgtgctgat tcccagatg caatcttccc ttctgcagca acacatggaa   | 1440 |
| ctcctgggtg cttcagtgat acccctgatg caccactgg cggacccag cactgtgttc    | 1500 |
| aaaaatggcg atgaggccga agactttgtg gaagttcacc tgcccgatgt acacgaaagg  | 1560 |
| atatctggag tagacctggg ccttcctaata tgggtaagt acgtgctcct gagtgcgggt  | 1620 |
| gccttgaccg ctttgatgct gatcattttt ctgatgacct gctggcggag ggtgaatcgc  | 1680 |
| tccgagccga cacagacaaa tctcagaggg acaggccggg aagtaagtgt gactccgcaa  | 1740 |
| tctggcaaga ttattagtag ttgggagagt tacaagtctg gaggagagac tgggttgaat  | 1800 |
| tttgatctgc tcaaacttgc aggcgatgta gaatcaaatc ctggaccgc cggggacagg   | 1860 |
| tccatagctc tcacgtttct cgcagttgga ggagttctgc tcttctctc cgtgaaactg   | 1920 |
| cacgtgaca ctgggtgtgc catagacatc agccggcaag agctgagatg tggaagtgga   | 1980 |
| gtgttcatac acaatgatgt                                              | 2000 |

&lt;210&gt; SEQ ID NO 60

&lt;211&gt; LENGTH: 2000

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 60

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| tcataagcg gacacactcg actgttttaa tcatacaaaa cactcctaata tgttgtaata  | 60  |
| tgtgtcacgc tcgacaaga atcgtgcttc tagagctaca gattctttgg tcctccggg    | 120 |
| ccgttctcgg cccgacagat aaacgatttt gcgccttacg gggcgcaaaa caggaaactaa | 180 |
| cctgaattcg ttttcttgcg tcccccttgc tgaccatata gacactagca aggagtccga  | 240 |
| gaaaacaaac atgggaacga ccataaaggg gaaacgaaac catttaaagg atagatatgg  | 300 |
| tagggactat tcgagcccgg aacctcaggg taactataag tggtaaacctc gacgggtttg | 360 |
| ttggagcagc aactcctact tcccactgta ttagaagagc ctaaaaggat gtacctcaac  | 420 |
| tttccccga tataaagtgc gtaattctac ttgcccgaat gaacatgtcc tcagcactgg   | 480 |
| cttcggctct gtatatgttt aaagcaccct atgcagtggt ggtggaagtt ctctttgtg   | 540 |
| aaggcgggtt gcggactgcg aacagcccgg cgaatgttga ccttctaccg tcctctagga  | 600 |
| gctatacttc ttagagacgt gttgggcata ggactaatgg taaccgagc ctgtcagttc   | 660 |

-continued

---

```

tgatggttcc tctcagaacca gtaatatagt ggttcgcacc ggctagaact aggaatacta 720
tctagggagc tgatcatcca aaaaggaccg cccttaacat cgccacaacg tcatagtcca 780
tggatgacga ggtgattggt gctgatgtga tatacctacg gactcctggg agctgagcca 840
tactcaacgc tgtaaaaatg cttgagtgcc cegttcgccc gtagattccc cagactttgt 900
acgccccaac aactactcgc ccccaacata ttagagaat ttccgcgac attcgacttt 960
gagacaccgc atgaccocga cgcggactac ctgcccgtga cccaccgata cgtctgttcg 1020
ttactttggt tcaccacagg gggaccagtc gaccaattag acgtgctgaa atccagactg 1080
ctttagctcg tggaaacacca cctccttgac cacttctttg cgcttctcac ggacctgct 1140
gaactctcat aactactggtg gtttaggcaa aggaagtctt ctgactcggg ggacctttc 1200
gaccacggtc ccaagccott ccgaatatga taaaagttgt tctgagaata cctccgcta 1260
cgggtaatat tcagtcaatc ctgaacctta ctctattaag ggaggtttcc tacagactct 1320
cagccacccct ctacgggtggg ggtacagtta cccacaaga aattgcetta gtaggaccct 1380
ggactgcctt tgcacgacta agggctctac gttagaaggg aagacgtcgt tgtgtacctt 1440
gaggaccaca gaagtcaacta tggggactac gtgggtgacc ggctggggtc gtgacacaag 1500
tttttaccgc tactccggct tctgaaacac cttcaagtgg acgggctaca tgtgctttcc 1560
tatagacctc atctggaccc ggaaggatta accccattca tgcacgagga ctcacgcccc 1620
cggaaactggc gaaactacga ctagtaaaaa gactactgga cgaccgcctc ccacttagcg 1680
aggctcggct gtgtcgtggt agagtctccc tgtccggccc ttcattcaca ctgaggcgtt 1740
agaccgttct aataatcctc aacctctca atgttcagac ctctctctg acccaactta 1800
aaactagacg agtttgaacg tccgctacat cttagttag gacctggcg ggccctgtcc 1860
aggatcagag agtgcaaaga gcgtcaacct cctcaagacg agaaggagag gcacttgac 1920
gtgagactgt gaccacacg gtatctgtag tcggccgttc tcgactctac accttcacct 1980
cacaagtatg tgttactaca 2000

```

&lt;210&gt; SEQ ID NO 61

&lt;211&gt; LENGTH: 1303

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 61

```

Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Asn Met
1 5 10 15
Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Gln
20 25 30
Lys Lys Arg Gly Gly Lys Thr Gly Ile Ala Val Ile Val Pro Gln Ala
35 40 45
Leu Leu Phe Val Pro Leu Leu Val Phe Pro Leu Cys Phe Gly Lys Phe
50 55 60
Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro Trp Ser Pro Ile Asp
65 70 75 80
Ile His His Leu Ser Cys Pro Asn Asn Leu Val Val Glu Asp Glu Gly
85 90 95
Cys Thr Asn Leu Ser Gly Phe Ser Tyr Met Glu Leu Lys Val Gly Tyr
100 105 110
Ile Ser Ala Ile Lys Met Asn Gly Phe Thr Cys Thr Gly Val Val Thr

```

-continued

| 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Ala | Glu | Thr | Tyr | Thr | Asn | Phe | Val | Gly | Tyr | Val | Thr | Thr | Thr | Phe |
| 130 |     |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |
| Lys | Arg | Lys | His | Phe | Arg | Pro | Thr | Pro | Asp | Ala | Cys | Arg | Ala | Ala | Tyr |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |
| Asn | Trp | Lys | Met | Ala | Gly | Asp | Pro | Arg | Tyr | Glu | Glu | Ser | Leu | His | Asn |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     |     | 175 |
| Pro | Tyr | Pro | Asp | Tyr | His | Trp | Leu | Arg | Thr | Val | Lys | Thr | Thr | Lys | Glu |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     |     | 190 |     |
| Ser | Leu | Val | Ile | Ile | Ser | Pro | Ser | Val | Ala | Asp | Leu | Asp | Pro | Tyr | Asp |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |
| Arg | Ser | Leu | His | Ser | Arg | Val | Phe | Pro | Gly | Gly | Asn | Cys | Ser | Gly | Val |
| 210 |     |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |
| Ala | Val | Ser | Ser | Thr | Tyr | Cys | Ser | Thr | Asn | His | Asp | Tyr | Thr | Ile | Trp |
| 225 |     |     |     |     | 230 |     |     |     |     |     | 235 |     |     |     | 240 |
| Met | Pro | Glu | Asn | Pro | Arg | Leu | Gly | Met | Ser | Cys | Asp | Ile | Phe | Thr | Asn |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     |     | 255 |
| Ser | Arg | Gly | Lys | Arg | Ala | Ser | Lys | Gly | Ser | Glu | Thr | Cys | Gly | Phe | Val |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     |     | 270 |     |
| Asp | Glu | Arg | Gly | Leu | Tyr | Lys | Ser | Leu | Lys | Gly | Ala | Cys | Lys | Leu | Lys |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     |     | 285 |     |     |
| Leu | Cys | Gly | Val | Leu | Gly | Leu | Arg | Leu | Met | Asp | Gly | Thr | Trp | Val | Ala |
| 290 |     |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |
| Met | Gln | Thr | Ser | Asn | Glu | Thr | Lys | Trp | Cys | Pro | Pro | Gly | Gln | Leu | Val |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |
| Asn | Leu | His | Asp | Phe | Arg | Ser | Asp | Glu | Ile | Glu | His | Leu | Val | Val | Glu |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     |     | 335 |
| Glu | Leu | Val | Lys | Lys | Arg | Glu | Glu | Cys | Leu | Asp | Ala | Leu | Glu | Ser | Ile |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     |     | 350 |     |
| Met | Thr | Thr | Lys | Ser | Val | Ser | Phe | Arg | Arg | Leu | Ser | His | Leu | Arg | Lys |
|     |     |     | 355 |     |     |     | 360 |     |     |     |     |     | 365 |     |     |
| Leu | Val | Pro | Gly | Phe | Gly | Lys | Ala | Tyr | Thr | Ile | Phe | Asn | Lys | Thr | Leu |
| 370 |     |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Met | Glu | Ala | Asp | Ala | His | Tyr | Lys | Ser | Val | Arg | Thr | Trp | Asn | Glu | Ile |
| 385 |     |     |     |     | 390 |     |     |     |     |     | 395 |     |     |     | 400 |
| Ile | Pro | Ser | Lys | Gly | Cys | Leu | Arg | Val | Gly | Gly | Arg | Cys | His | Pro | His |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     |     | 415 |
| Val | Asn | Gly | Val | Phe | Phe | Asn | Gly | Ile | Ile | Leu | Gly | Pro | Asp | Gly | Asn |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     |     | 430 |     |
| Val | Leu | Ile | Pro | Glu | Met | Gln | Ser | Ser | Leu | Leu | Gln | Gln | His | Met | Glu |
|     |     |     | 435 |     |     |     | 440 |     |     |     |     |     | 445 |     |     |
| Leu | Leu | Val | Ser | Ser | Val | Ile | Pro | Leu | Met | His | Pro | Leu | Ala | Asp | Pro |
| 450 |     |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Ser | Thr | Val | Phe | Lys | Asn | Gly | Asp | Glu | Ala | Glu | Asp | Phe | Val | Glu | Val |
| 465 |     |     |     |     | 470 |     |     |     |     |     | 475 |     |     |     | 480 |
| His | Leu | Pro | Asp | Val | His | Glu | Arg | Ile | Ser | Gly | Val | Asp | Leu | Gly | Leu |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     |     | 495 |
| Pro | Asn | Trp | Gly | Lys | Tyr | Val | Leu | Leu | Ser | Ala | Gly | Ala | Leu | Thr | Ala |
|     |     |     | 500 |     |     |     |     |     | 505 |     |     |     |     |     | 510 |
| Leu | Met | Leu | Ile | Ile | Phe | Leu | Met | Thr | Cys | Trp | Arg | Arg | Val | Asn | Arg |
|     |     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |
| Ser | Glu | Pro | Thr | Gln | His | Asn | Leu | Arg | Gly | Thr | Gly | Arg | Glu | Val | Ser |
| 530 |     |     |     |     |     | 535 |     |     |     |     |     |     | 540 |     |     |

-continued

---

Val Thr Pro Gln Ser Gly Lys Ile Ile Ser Ser Trp Glu Ser Tyr Lys  
 545 550 555 560  
 Ser Gly Gly Glu Thr Gly Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly  
 565 570 575  
 Asp Val Glu Ser Asn Pro Gly Pro Gly Gly Lys Thr Gly Ile Ala Val  
 580 585 590  
 Met Ile Gly Leu Ile Ala Cys Val Gly Ala Val Thr Leu Ser Asn Phe  
 595 600 605  
 Gln Gly Lys Val Met Met Thr Val Asn Ala Thr Asp Val Thr Asp Val  
 610 615 620  
 Ile Thr Ile Pro Thr Ala Ala Gly Lys Asn Leu Cys Ile Val Arg Ala  
 625 630 635 640  
 Met Asp Val Gly Tyr Met Cys Asp Asp Thr Ile Thr Tyr Glu Cys Pro  
 645 650 655  
 Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys Thr  
 660 665 670  
 Lys Ser Ala Val Tyr Val Arg Tyr Gly Arg Cys Thr Lys Thr Arg His  
 675 680 685  
 Ser Arg Arg Ser Arg Arg Ser Leu Thr Val Gln Thr His Gly Glu Ser  
 690 695 700  
 Thr Leu Ala Asn Lys Lys Gly Ala Trp Met Asp Ser Thr Lys Ala Thr  
 705 710 715 720  
 Arg Tyr Leu Val Lys Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly Tyr  
 725 730 735  
 Ala Leu Val Ala Ala Val Ile Gly Trp Met Leu Gly Ser Asn Thr Met  
 740 745 750  
 Gln Arg Val Val Phe Val Val Leu Leu Leu Leu Val Ala Pro Ala Tyr  
 755 760 765  
 Ser Phe Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly Val  
 770 775 780  
 Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys Val  
 785 790 795 800  
 Thr Ile Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met Asn  
 805 810 815  
 Met Glu Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu Ala  
 820 825 830  
 Thr Val Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Ala Met Gly Glu  
 835 840 845  
 Ala His Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln Gly  
 850 855 860  
 Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly  
 865 870 875 880  
 Ser Ile Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile Gly  
 885 890 895  
 Arg Thr Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val  
 900 905 910  
 His Gly Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln Val  
 915 920 925  
 Gly Ala Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro Ser  
 930 935 940  
 Tyr Thr Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys Glu  
 945 950 955 960

-continued

---

Pro Arg Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val Gly  
                                   965                                  970                                  975  
 Thr Lys Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn Leu  
                                   980                                  985                                  990  
 Pro Trp Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr Leu  
                                   995                                  1000                                  1005  
 Met Glu Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala  
   1010                                  1015                                  1020  
 Leu Gly Ser Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala  
   1025                                  1030  
 Ile Pro Val Glu Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly  
   1040                                  1045                                  1050  
 His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly  
   1055                                  1060                                  1065  
 Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr  
   1070                                  1075                                  1080  
 Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr  
   1085                                  1090                                  1095  
 Thr Gly Thr Asp Gly Pro Cys Lys Val Pro Ile Ser Ser Val Ala  
   1100                                  1105                                  1110  
 Ser Leu Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val Asn  
   1115                                  1120                                  1125  
 Pro Phe Val Ser Val Ala Thr Ala Asn Ala Lys Val Leu Ile Glu  
   1130                                  1135                                  1140  
 Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly  
   1145                                  1150                                  1155  
 Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly Ser Ser Ile  
   1160                                  1165                                  1170  
 Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala  
   1175                                  1180                                  1185  
 Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly Gly Val  
   1190                                  1195                                  1200  
 Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe Gly Gly Ala  
   1205                                  1210                                  1215  
 Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu  
   1220                                  1225                                  1230  
 Leu Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg  
   1235                                  1240                                  1245  
 Ser Ile Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe  
   1250                                  1255                                  1260  
 Leu Ser Val Asn Val Glu His Ala Asp Thr Gly Cys Ala Ile Asp  
   1265                                  1270                                  1275  
 Ile Ser Arg Gln Glu Leu Arg Cys Gly Ser Gly Val Phe Ile His  
   1280                                  1285                                  1290  
 Asn Asp Val Glu Ala Trp Met Asp Arg Tyr  
   1295                                  1300

&lt;210&gt; SEQ ID NO 62

&lt;211&gt; LENGTH: 4000

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 62

-continued

---

|            |             |            |            |            |             |      |
|------------|-------------|------------|------------|------------|-------------|------|
| agtagttcgc | ctgtgtgagc  | tgacaaactt | agtagtgttt | gtgaggatta | acaacaatta  | 60   |
| acacagtgcg | agctgtttct  | tagcacgaag | atctcgatgt | ctaagaaacc | aggagggccc  | 120  |
| ggcaagagcc | gggctgtcaa  | tatgctaaaa | cgcggaatgc | cccgcgtggt | gtccttgatt  | 180  |
| ggacttaagc | aaaagaagcg  | agggggcaag | actggtatag | ctgtgatcgt | tcctcaggct  | 240  |
| cttttgtttg | taccttgct   | ggtatttccc | ctttgctttg | gtaaatttcc | tatctatacc  | 300  |
| atccctgata | agctcggggc  | ttggagtcce | attgatattc | accatttgag | ctgccc aaac | 360  |
| aacctcgtcg | ttgaggatga  | agggtgcact | aatctttctg | gattttccta | catggagttg  | 420  |
| aaagtgggct | atatttcagc  | cattaagatg | aacggcttta | cttgtagagg | agtcgtgacc  | 480  |
| gaagccgaga | catatacaaa  | ttctgtggga | tacgtcacca | ccaccttcaa | gagaaaaac   | 540  |
| ttccgcccaa | cgctgacgc   | ttgtcggggc | gcttacaact | ggaagatggc | aggagatcct  | 600  |
| cgatatgaag | aatctctgca  | caaccctgat | cctgattacc | attggctgcg | gacagtcaag  | 660  |
| actaccaagg | agagtctggt  | cattatatca | ccaagcgtgg | ccgatcttga | tccttatgat  | 720  |
| agatccctgc | acagtagggt  | ttttcctggc | gggaattgta | gcggtgttgc | agtatcaagt  | 780  |
| acctactgct | ccactaacca  | cgactacact | atatggatgc | ctgagaacct | tcgactcggg  | 840  |
| atgagttgcg | acatttttac  | gaactcacgg | ggcaagcggg | catctaaggg | gtctgaaaca  | 900  |
| tgccggtttg | ttgatgagcg  | ggggttgat  | aaatctctta | aaggcgcctg | taagctgaaa  | 960  |
| ctctgtggcg | tactggggct  | gcccctgatg | gacggcacat | gggtggctat | gcagacaagc  | 1020 |
| aatgaaacaa | agtggtgtcc  | cctcggtcag | ctggttaatc | tgcacgactt | taggtctgac  | 1080 |
| gaaatcgagc | accttgtggt  | ggaggaactg | gtgaagaaac | gcgaagagtg | cctggacgca  | 1140 |
| cttgagagta | ttatgaccac  | caaatcogtt | tccttcagaa | gactgagcca | cctgcgaaag  | 1200 |
| ctggtgccag | gggtcgggaa  | ggcttatact | attttcaaca | agactcttat | ggaggcggat  | 1260 |
| gcccattata | agtcagttag  | gacttggaat | gagataattc | cctccaaagg | atgtctgaga  | 1320 |
| gtcggtgggg | gatgcccccc  | ccatgtcaat | ggggtgttct | ttaacggaat | catcctggga  | 1380 |
| cctgacggga | acgtgctgat  | tcccagatg  | caatcttccc | ttctgcagca | acacatggaa  | 1440 |
| ctcctggtgt | cttcagtgat  | accctgatg  | caccactgg  | ccgaccccag | cactgtgttc  | 1500 |
| aaaaatggcg | atgaggccga  | agactttgtg | gaagttcacc | tgcccgatgt | acacgaaagg  | 1560 |
| atatctggag | tagacctggg  | ccttctaat  | tgggtaagt  | acgtgctcct | gagtgcgggt  | 1620 |
| gccttgaccg | ctttgatgct  | gatcattttt | ctgatgacct | gctggcggag | ggtgaatcgc  | 1680 |
| tccgagccga | cacagcacia  | tctcagaggg | acaggccggg | aagtaagtgt | gactccgcaa  | 1740 |
| tctggcaaga | ttattagtag  | ttgggagagt | tacaagtctg | gaggagagac | tgggttgaat  | 1800 |
| tttgatctgc | tcaaacttgc  | aggcagtgta | gaatcaaatc | ctggacccgg | aggaaagacc  | 1860 |
| ggatttgca  | tcattgattg  | cctgatcgcc | tgcgtaggag | cagttaccct | ctctaacttc  | 1920 |
| caagggaagg | tgatgatgac  | ggtaaatgct | actgaagtca | cagatgtcat | cacgattcca  | 1980 |
| acagctgctg | gaaagaacct  | atgcattgtc | agagcaatgg | atgtgggata | catgtgcgat  | 2040 |
| gatactatca | cttatgaatg  | cccagtgctg | tcggctggta | atgatccaga | agacatcgac  | 2100 |
| tgttggtgca | caaagtacgc  | agtctacgtc | aggtaggaa  | gatgcaccaa | gacacgccac  | 2160 |
| tcaagacgca | gtcggagggtc | actgacagtg | cagacacacg | gagaaagcac | tctagcgaac  | 2220 |
| aagaaggggg | cttgatgga   | cagcaccaag | gccacaaggt | atttggtaaa | aacagaatca  | 2280 |
| tggatcttga | ggaacctg    | atatgccctg | gtggcagccg | tcattggttg | gatgcttggg  | 2340 |
| agcaacacca | tgcagagagt  | tgtgtttgtc | gtgctattgc | ttttggtggc | cccagcttac  | 2400 |

-continued

---

```

agctttaact gccttgaat gagcaacaga gacttcttgg aaggagtgc tggagcaaca 2460
tgggttgatt tggttctoga aggcgacagc tgcgtgacta tcatgtctaa ggacaagcct 2520
accatcgatg tgaagatgat gaatatggag gcggccaacc tggcagaggt ccgcagttat 2580
tgctatattg ctaccgctag cgatctctcc accaaagctg cgtgcccggc catgggagaa 2640
gctcacaatg acaaacgtgc tgaccagct tttgtgtgca gacaaggagt ggtggacagg 2700
ggctggggca acggctcggg actatattggc aaaggaagca ttgacacatg cgccaaattt 2760
gectgctcta ccaaggcaat aggaagaacc attttgaag agaatatcaa gtacgaagtg 2820
gccatttttg tccatggacc aactactgtg gagtcgcacg gaaactactc cacacaggtt 2880
ggagccactc aggcaggagg attcagcatc actcctcggg cgccttcata cacactaaag 2940
cttgagaaat atggagaggt gacagtggac tgtgaaccac ggtcagggat tgacaccaat 3000
gcatactacg tgatgactgt tggaaacaaag acgttcttgg tccatcgtga gtggttcatg 3060
gacctcaacc tcccttgagg cagtgtctga agtactgtgt ggaggaacag agagacgtta 3120
atggagtttg aggaaccaca cgccacgaag cagtctgtga tagcattggg ctcaacagag 3180
ggagctctgc atcaagcttt ggctggagcc attcctgtgg aattttcaag caacactgtc 3240
aagttgacgt cgggtcattt gaagtgtaga gtgaagatgg aaaaattgca gttgaagggg 3300
acaacctatg gcgtctgttc aaaggcttcc aagtttcttg ggactcccgc agacacaggt 3360
cacggcactg tgggtttgga attgcagtac actggcacgg atggaccttg caaagttcct 3420
atctcgtcag tggcttcatt gaacgacctc acgccagtg gcagattggt cactgtcaac 3480
ccttttgttt cagtggccac ggccaacgct aaggtcctga ttgaattgga accacccttt 3540
ggagactcat acatagtggg gggcagagga gaacaacaga tcaatcacca ctggcacaag 3600
tctggaagca gcattggcaa agcctttaca accaccctca aaggagcgca gagactagcc 3660
gctctaggag acacagcttg ggactttgga tcagttggag ggggtgtcac ctacagttggg 3720
aaggctgtcc atcaagtgtt cggaggagca ttccgctcac tgttcggagg catgtcctgg 3780
ataacgcaag gattgctggg ggctctcctg ttgtggatgg gcatcaatgc tcgtgacagg 3840
tccatagctc tcacgtttct cgcagttgga ggagttctgc tcttctctc cgtgaacgtg 3900
cacgctgaca ctgggtgtgc catagacatc agccggcaag agctgagatg tggaaagtga 3960
gtgttcatac acaatgatgt ggaggcttgg atggaccggt 4000

```

&lt;210&gt; SEQ ID NO 63

&lt;211&gt; LENGTH: 4000

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 63

```

tcatcaagcg gacacactcg actgttttaa tcatcacaaa cactccta at tgtgtttaat 60
tgtgtcacgc tcgacaaaga atcgtgcttc tagagctaca gattctttgg tctcctccgg 120
ccgttctcgg cccgacagtt atacgatttt gcgccttacg gggcgcaaa caggaactaa 180
cctgaattcg ttttcttcgc tcccccttc tgaccatata gacactagca aggagtccga 240
gaaaacaac atgggaacga ccataaagg gaaacgaaac catttaaagg atagatatgg 300
tagggactat tcgagcccg aacctcagg taactataag tggtaaactc gacgggtttg 360
ttggagcagc aactcctact tcccactgga ttagaaagac ctaaaaggat gtacctcaac 420

```

-continued

---

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tffcaccoga  | tataaagtcg  | gtaattctac  | ttgccgaaat  | gaacatgtcc  | tcagcactgg  | 480  |
| cttcggctct  | gtatatgttt  | aaagcacccct | atgcagtggt  | ggtggaagtt  | ctcttttgtg  | 540  |
| aaggcgggtt  | gcggaactgcg | aacagcccgg  | cgaatgttga  | cctctaccg   | tcctctagga  | 600  |
| gctatacttc  | ttagagacgt  | ggtgggcata  | ggactaatgg  | taaccgacgc  | ctgtcagttc  | 660  |
| tgatggttcc  | tctcagacca  | gtaatatagt  | ggttcgcacc  | ggctagaact  | aggaatacta  | 720  |
| tctagggacg  | tgatcatcca  | aaaaggaccg  | cccttaacat  | cgccacaacg  | tcatagtcca  | 780  |
| tggatgacga  | ggtgatgggt  | gctgatgtga  | tatacctacg  | gactcctggg  | agctgagcca  | 840  |
| tactcaacgc  | tgtaaaaatg  | cttgagtgcc  | ccgttcgccc  | gtagattccc  | cagactttgt  | 900  |
| acgcccacaac | aactactcgc  | ccccaacata  | tttagagaat  | ttccgcggaac | attcagacttt | 960  |
| gagacaccgc  | atgaccocga  | cgcggactac  | ctgccgtgta  | cccaccgata  | cgtctgttcg  | 1020 |
| ttactttgtt  | tcaccacagg  | gggaccagtc  | gaccaattag  | acgtgctgaa  | atccagactg  | 1080 |
| ctttagctcg  | tggaacacca  | cctccttgac  | cactctcttg  | cgtctctcac  | ggacctgcgt  | 1140 |
| gaactctcat  | aatactgggtg | gtttaggcaa  | aggaagtctt  | ctgactcggg  | ggacctttc   | 1200 |
| gaccacggtc  | ccaagccctt  | ccgaatatga  | taaaagttgt  | tctgagaata  | cctccgccta  | 1260 |
| cgggtaatat  | tcagtcaatc  | ctgaacctta  | ctctattaag  | ggaggtttcc  | tacagactct  | 1320 |
| cagccaccct  | ctacgggtggg | ggtacagtta  | ccccacaaga  | aattgcctta  | gtaggaccct  | 1380 |
| ggactgcctt  | tgcacgacta  | agggtctctac | gttagaaggg  | aagacgtcgt  | tgtgtacctt  | 1440 |
| gaggaccaca  | gaagtcaacta | tggggactac  | gtgggtgacc  | ggctggggtc  | gtgacacaag  | 1500 |
| ttttaccgc   | tactccggct  | tctgaaacac  | cttcaagtgg  | acgggttaca  | tgtgctttcc  | 1560 |
| tatagaccct  | atctggaccc  | ggaaggatta  | acccattcca  | tgcacgagga  | ctcacgccc   | 1620 |
| cggaaactggc | gaaactacga  | ctagtaaaaa  | gactactgga  | cgaccgcctc  | ccacttagcg  | 1680 |
| aggctcggct  | gtgtcgtggt  | agagtctccc  | tgtccggccc  | ttcattcaca  | ctgaggcgtt  | 1740 |
| agaccgttct  | aataatcctc  | aaccctctca  | atgttcagac  | ctcctctctg  | acccaactta  | 1800 |
| aaactagaag  | agtttgaaag  | tcgcctacat  | cttagtttag  | gacctgggac  | tcctttctgg  | 1860 |
| ccataacgct  | agtactaacc  | ggactagcgg  | acgcctcctc  | gtcaatggga  | gagattgaag  | 1920 |
| gttcctctcc  | actactactg  | ccatttacga  | tgaactgcagt | gtctacagta  | gtgctaagggt | 1980 |
| tgctgacgac  | ctttcttgga  | tacgtaacag  | tctcgttacc  | tacaccctat  | gtacacgcta  | 2040 |
| ctatgatagt  | gaatacttac  | gggtcacgac  | agccgaccat  | tactaggtct  | tctgtagctg  | 2100 |
| acaaccacgt  | gtttcagtcg  | tcagatgcag  | tccatacctt  | ctacgtgggt  | ctgtgcgggtg | 2160 |
| agttctgcgt  | cagcctccag  | tgaactgtcac | gtctgtgtgc  | ctctttcgtg  | agatcgtttg  | 2220 |
| ttcttcccc   | gaacctacct  | gtcgtgggtc  | cgggtgtcca  | taaaccattt  | ttgtcttagt  | 2280 |
| acctagaact  | ccttgggacc  | tatacgggac  | caccgtcggc  | agtaaccaac  | ctacgaacct  | 2340 |
| tcgttggtgt  | acgtctctca  | acacaaacag  | cacgataacg  | aaaaccaccg  | gggtcgaatg  | 2400 |
| tcgaaattga  | cggaaacctta | ctcgttgtct  | ctgaagaacc  | ttcctcacag  | acctcgttgt  | 2460 |
| acccacctaa  | accaagagct  | tcgcctgtcg  | acgcactgat  | agtacagatt  | cctgttcogga | 2520 |
| tggtagctac  | acttctacta  | cttatacctc  | cgccgggttg  | accgtctcca  | ggcgtcaata  | 2580 |
| acgataaacc  | gatggcagtc  | gctagagagg  | tggtttcgac  | gcacgggccc  | gtaccctctt  | 2640 |
| cgagtgttac  | tgtttgcaag  | actgggtcga  | aaacacacgt  | ctgttcctca  | ccacctgtcc  | 2700 |
| ccgacccctg  | tgccgacgcc  | tgataaaccc  | ttctctctgt  | aactgtgtac  | gcggtttaaa  | 2760 |
| cggacgagat  | ggttccgtta  | tccttctgtg  | taaaactttc  | tcttatagtt  | catgcttcac  | 2820 |

-continued

---

```

cggtaaaaac aggtacctgg ttgatgacac ctcagcgtgc ctttgatgag gtgtgtocaa 2880
cctcgggtgag tccgtccctc taagtcgtag tgaggacgcc gcggaagtat gtgtgatttc 2940
gaacctctta tacctctcca ctgtcaacctg acacttggtg ccagtcctca actgtgggta 3000
cgtatgatgc actactgaca accttgtttc tgcaagaacc aggtagcact caccaagtac 3060
ctggagttgg agggaacctc gtcacgacct tcatgacaca cctccttgtc tctctgcaat 3120
tacctcaaac tccttggtgt gcggtgcttc gtcagacct atcgtaacct gagtgttctc 3180
cctcgagacg tagttcgaaa ccgacctcgg taaggacacc ttaaaagttc gttgtgacag 3240
ttcaactgca gcccagtaaa cttcacatct cacttctacc tttttaacgt caacttcct 3300
tgttggatag cgcagacaag tttccgaaaag ttcaaagaac cctgaggggcg tctgtgtcca 3360
gtgccgtgac accacaacct taacgtcatg tgacctgcc tacctggaac gtttcaagga 3420
tagagcagtc accgaagtaa cttgctggat tgcggtcacc cgtctaacca gtgacagttg 3480
ggaaaacaaa gtcaccgggtg ccggttgcca ttccaggact aacttaacct tgggtgggaaa 3540
cctctgagta tgatcaacca ccgctctcct cttgttgtct agttagtggt gacctgttc 3600
agaccttctg cgtaacctgt tcgaaaatgt tgggtggagt ttcctcgcgt ctctgatcgg 3660
cgagatcctc tgtgtcgaac cctgaaacct agtcaacctc cccacaagtg gagtcaacct 3720
ttccgacagg tagttcaaaa gctcctcctg aaggcagtg acaagcctcc gtacaggacc 3780
tattgcgttc ctaacgacct ccgagaggac aacacctacc cgtagttacg agcactgtcc 3840
aggtatcgag agtgcaaaga gcgtcaacct cctcaagacg agaaggagag gcacttgcac 3900
gtgagactgt gaccacacag gtatctgtag tcggccgttc tcgactctac accttcacct 3960
cacaagtatg tgttactaca cctccgaacc tacctggcca 4000

```

&lt;210&gt; SEQ ID NO 64

&lt;211&gt; LENGTH: 702

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 64

```

Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Tyr Leu
1 5 10 15
Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Arg
20 25 30
Ala Met Leu Ser Leu Ile Asp Gly Lys Gly Pro Ile Arg Phe Val Leu
35 40 45
Ala Leu Leu Ala Phe Phe Arg Phe Thr Ala Ile Ala Pro Thr Arg Ala
50 55 60
Val Leu Asp Arg Trp Arg Gly Val Asn Lys Gln Thr Ala Met Lys His
65 70 75 80
Leu Leu Ser Phe Lys Lys Glu Leu Gly Thr Leu Thr Ser Ala Ile Asn
85 90 95
Arg Arg Ser Ser Lys Gln Lys Lys Arg Gly Gly Lys Thr Gly Ile Ala
100 105 110
Val Ile Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Pro
115 120 125
Leu Cys Phe Gly Lys Phe Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly
130 135 140
Pro Trp Ser Pro Ile Asp Ile His His Leu Ser Cys Pro Asn Asn Leu

```

-continued

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 145 |     |     |     | 150 |     |     |     |     |     | 155 |     |     |     | 160 |     |
| Val | Val | Glu | Asp | Glu | Gly | Cys | Thr | Asn | Leu | Ser | Gly | Phe | Ser | Tyr | Met |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |
| Glu | Leu | Lys | Val | Gly | Tyr | Ile | Ser | Ala | Ile | Lys | Met | Asn | Gly | Phe | Thr |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |
| Cys | Thr | Gly | Val | Val | Thr | Glu | Ala | Glu | Thr | Tyr | Thr | Asn | Phe | Val | Gly |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |
| Tyr | Val | Thr | Thr | Thr | Phe | Lys | Arg | Lys | His | Phe | Arg | Pro | Thr | Pro | Asp |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |
| Ala | Cys | Arg | Ala | Ala | Tyr | Asn | Trp | Lys | Met | Ala | Gly | Asp | Pro | Arg | Tyr |
|     | 225 |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |
| Glu | Glu | Ser | Leu | His | Asn | Pro | Tyr | Pro | Asp | Tyr | His | Trp | Leu | Arg | Thr |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |
| Val | Lys | Thr | Thr | Lys | Glu | Ser | Leu | Val | Ile | Ile | Ser | Pro | Ser | Val | Ala |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     |     | 270 |     |
| Asp | Leu | Asp | Pro | Tyr | Asp | Arg | Ser | Leu | His | Ser | Arg | Val | Phe | Pro | Gly |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |
| Gly | Asn | Cys | Ser | Gly | Val | Ala | Val | Ser | Ser | Thr | Tyr | Cys | Ser | Thr | Asn |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |
| His | Asp | Tyr | Thr | Ile | Trp | Met | Pro | Glu | Asn | Pro | Arg | Leu | Gly | Met | Ser |
|     | 305 |     |     |     |     | 310 |     |     |     | 315 |     |     |     |     | 320 |
| Cys | Asp | Ile | Phe | Thr | Asn | Ser | Arg | Gly | Lys | Arg | Ala | Ser | Lys | Gly | Ser |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |
| Glu | Thr | Cys | Gly | Phe | Val | Asp | Glu | Arg | Gly | Leu | Tyr | Lys | Ser | Leu | Lys |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |
| Gly | Ala | Cys | Lys | Leu | Lys | Leu | Cys | Gly | Val | Leu | Gly | Leu | Arg | Leu | Met |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
| Asp | Gly | Thr | Trp | Val | Ala | Met | Gln | Thr | Ser | Asn | Glu | Thr | Lys | Trp | Cys |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Pro | Pro | Gly | Gln | Leu | Val | Asn | Leu | His | Asp | Phe | Arg | Ser | Asp | Glu | Ile |
|     | 385 |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |
| Glu | His | Leu | Val | Val | Glu | Glu | Leu | Val | Lys | Lys | Arg | Glu | Glu | Cys | Leu |
|     |     |     | 405 |     |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Asp | Ala | Leu | Glu | Ser | Ile | Met | Thr | Thr | Lys | Ser | Val | Ser | Phe | Arg | Arg |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     |     | 430 |     |
| Leu | Ser | His | Leu | Arg | Lys | Leu | Val | Pro | Gly | Phe | Gly | Lys | Ala | Tyr | Thr |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     |     | 445 |     |     |
| Ile | Phe | Asn | Lys | Thr | Leu | Met | Glu | Ala | Asp | Ala | His | Tyr | Lys | Ser | Val |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Arg | Thr | Trp | Asn | Glu | Ile | Ile | Pro | Ser | Lys | Gly | Cys | Leu | Arg | Val | Gly |
|     | 465 |     |     | 470 |     |     |     |     |     | 475 |     |     |     |     | 480 |
| Gly | Arg | Cys | His | Pro | His | Val | Asn | Gly | Val | Phe | Phe | Asn | Gly | Ile | Ile |
|     |     |     | 485 |     |     |     |     |     | 490 |     |     |     |     | 495 |     |
| Leu | Gly | Pro | Asp | Gly | Asn | Val | Leu | Ile | Pro | Glu | Met | Gln | Ser | Ser | Leu |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     |     | 510 |     |
| Leu | Gln | Gln | His | Met | Glu | Leu | Leu | Val | Ser | Ser | Val | Ile | Pro | Leu | Met |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     |     | 525 |     |     |
| His | Pro | Leu | Ala | Asp | Pro | Ser | Thr | Val | Phe | Lys | Asn | Gly | Asp | Glu | Ala |
|     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Glu | Asp | Phe | Val | Glu | Val | His | Leu | Pro | Asp | Val | His | Glu | Arg | Ile | Ser |
|     | 545 |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Gly | Val | Asp | Leu | Gly | Leu | Pro | Asn | Trp | Gly | Lys | Tyr | Val | Leu | Leu | Ser |
|     |     |     | 565 |     |     |     |     |     | 570 |     |     |     |     | 575 |     |

-continued

Ala Gly Ala Leu Thr Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys  
 580 585 590

Trp Arg Arg Val Asn Arg Ser Glu Pro Thr Gln His Asn Leu Arg Gly  
 595 600 605

Thr Gly Arg Glu Val Ser Val Thr Pro Gln Ser Gly Lys Ile Ile Ser  
 610 615 620

Ser Trp Glu Ser Tyr Lys Ser Gly Gly Glu Thr Gly Leu Asn Phe Asp  
 625 630 635 640

Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly Pro Ala Arg  
 645 650 655

Asp Arg Ser Ile Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu  
 660 665 670

Phe Leu Ser Val Asn Val His Ala Asp Thr Gly Cys Ala Ile Asp Ile  
 675 680 685

Ser Arg Gln Glu Leu Arg Cys Gly Ser Gly Val Phe Ile His  
 690 695 700

&lt;210&gt; SEQ ID NO 65

&lt;211&gt; LENGTH: 2200

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 65

```

agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta 60
acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc 120
ggcaagagcc gggctgtcta tttgctaaaa cgcggaatgc cccgcgtggt gtccttgatt 180
ggacttaaga gggctatggt gagcctgata gacggcaagg ggccaatacg atttgtgttg 240
gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga 300
tggagagggt tgaacaaaca aacagcgatg aaacaccttc tgagtttcaa gaaggaacta 360
gggaccttga ccagtgtatc caatcggcgg agctcaaagc aaaagaagcg agggggcaag 420
actggtatag ctgtgatcgt tctcaggct cttttgtttg tacccttgcg ggtatattccc 480
ctttgcttgg gtaaatttcc tatctatacc atccctgata agctcggggc ttggagtccc 540
attgatattc accatttgag ctgcccacaa aacctcgtcg ttgaggatga aggggtcact 600
aatctttctg gattttccta catggagttg aaagtgggct atatttcagc cattaagatg 660
aacggcttta cttgtacagg agtcgtgacc gaagccgaga catatacaaa tttcgtggga 720
tacgtcacca ccacctcaa gagaaaacac ttccgcccac cgctgacgc ttgtcggggc 780
gcttacaact ggaagatggc aggagatcct cgatatgaag aatctctgca caaccggtat 840
cctgattacc attggctgcg gacagtcaag actaccaagg agagtctggt cattatatca 900
ccaagcgtgg ccgatcttga tccttatgat agatccctgc acagtagggg ttttcctggc 960
gggaattgta gcggtgttgc agtatcaagt acctactgct ccaactaacca cgactacact 1020
atatggatgc ctgagaaccc tcgactcggg atgagttgcy acatttttac gaactcacgg 1080
ggcaagcggg catctaaggg gtctgaaaca tgccgggtttg ttgatgagcg ggggttgat 1140
aaatctctta aaggcgcctg taagctgaaa ctctgtggcg tactggggct gcgcctgatg 1200
gacggcacat ggggtgctat gcagacaagc aatgaaacaa agtgggtgcc ccctggtcag 1260
ctggttaatc tgcacgactt taggtctgac gaaatcgagc accttgtggg ggaggaactg 1320

```

-continued

---

|                                                                   |      |
|-------------------------------------------------------------------|------|
| gtgaagaaac gcgaagagtg cctggacgca cttgagagta ttatgaccac caaatccggt | 1380 |
| tccttcagaa gactgagcca cctgcaaaag ctgggtgccag gggtcggaa ggcttatact | 1440 |
| atthtcaaca agactcttat ggaggcggat gccattata agtcagttag gacttggaat  | 1500 |
| gagataatc cctccaaagg atgtctgaga gtcggtggga gatgccacc ccatgtcaat   | 1560 |
| ggggtgttct ttaacggaat catcctggga cctgacggga acgtgctgat tcccagatg  | 1620 |
| caatcttccc ttctgcagca acacatggaa ctctgggtgt cttcagtgat acccctgatg | 1680 |
| caccactgg cgcacccag cactgtgttc aaaaatggcg atgaggccga agactttgtg   | 1740 |
| gaagttcacc tgcccgatgt acacgaaagg atatctggag tagacctggg ccttcctaat | 1800 |
| tgggtaagt acgtgctcct gactgctggg gccttgaccg ctttgatgct gatcattttt  | 1860 |
| ctgatgacct gctggcggag ggtgaatcgc tccgagccga cacagcaca tctcagagg   | 1920 |
| acaggccggg aagtaagtgt gactccgcaa tctggcaaga ttattagtag ttgggagagt | 1980 |
| tacaagtctg gaggagagac tgggttgaat tttgatctgc tcaaactgc agcgatgta   | 2040 |
| gaatcaaatc ctggaccgc cgggacagg tccatagctc tcacgtttct cgcagttgga   | 2100 |
| ggagttctgc tcttctctc cgtgaacgtg cacgctgaca ctgggtgtgc catagacatc  | 2160 |
| agccggcaag agctgagatg tggaaagtga gtgttcatac                       | 2200 |

&lt;210&gt; SEQ ID NO 66

&lt;211&gt; LENGTH: 2200

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 66

|                                                                   |      |
|-------------------------------------------------------------------|------|
| tcataagcg gacacactcg actgttttaa tcatacaaaa cactcctaat tgttgtaaat  | 60   |
| tgtgtcacgc tcgacaaaga atcgtgcttc tagagctaca gattctttgg tcctcccggg | 120  |
| ccgttctcgg cccgacagat aaacgatttt gcgccttacg gggcgcaaaa caggaactaa | 180  |
| cctgaattct cccgatacaa ctccgactag ctgcccgttc ccggttatgc taaacacaac | 240  |
| cgagagaacc gcaagaagtc caagtgtcgt taacgaggct gggctcgtca cgacctagct | 300  |
| acctctccac acttgtttgt ttgtcgtcct tttgtggaag actcaaatgt cttccttgat | 360  |
| ccctggaact ggtcacgata gttagccgcc tcgagtttcg ttttcttcgc tccccggtc  | 420  |
| tgaccatac gacactagca aggagtccga gaaaacaaac atgggaacga ccataaagg   | 480  |
| gaaacgaaac catttaaagg atagatatgg tagggactat tcgagcccg aaacctcagg  | 540  |
| taactataag tggtaaacct gacgggtttg ttggagcagc aactcctact tcccactgta | 600  |
| ttagaaagac ctaaaaggat gtacctcaac tttcaccga tataaagtcg gtaattctac  | 660  |
| tgccgaaat gaacatgtcc tcagcactgg ctccggctct gtatatgttt aaagcaccct  | 720  |
| atgcagtggg ggtggaagtt ctctttttgt aaggcgggtt gcggactgcg aacagcccg  | 780  |
| cgaatgttga ccttctaacc tcctctagga gctatacttc ttagagacgt gttgggcata | 840  |
| ggactaatgg taaccgacgc ctgtcagttc tgatggttcc tctcagacca gtaatatagt | 900  |
| ggttcgcacc ggctagaact aggaatacta tctagggacg tgcacatcca aaaaggaccg | 960  |
| cccttaacat cgcacaacg tcatagttca tggatgacga ggtgattggt gctgatgtga  | 1020 |
| tatacctacg gactcttggg agctgagcca tactcaacgc tgtaaaaatg cttgagtgcc | 1080 |
| ccgttcgccc gtagattccc cagactttgt acgccccaac aactactcgc ccccaacata | 1140 |
| tttagagaat ttccgcgac attcgaactt gagacaccgc atgacccga cgcggactac   | 1200 |

-continued

```

ctgccgtgta cccaccgata cgtctgttcg ttactttggt tcaccacagg gggaccagtc 1260
gaccaattag acgtgctgaa atccagactg ctttagctcg tggaacacca cctccttgac 1320
cacttctttg cgcttctcac ggacctgctg gaactctcat aatactggtg gtttaggcaa 1380
aggaagtctt ctgactcggg ggacgcttcc gaccacggtc ccaagccctt cgaatatga 1440
taaaagtgtg tctgagaata cctccgcta cgggtaatat tcagtcaatc ctgaacctta 1500
ctctattaag ggaggtttcc tacagactct cagccaccct ctacgggtggg ggtacagtta 1560
ccccacaaga aattgcctta gtaggacctt ggactgccct tgcacgacta agggctctac 1620
gttagaaggg aagacgtcgt tgtgtacctt gaggaccaca gaagtcacta tggggactac 1680
gtgggtgacc ggctggggtc gtgacacaag tttttaccgc tactccggct tctgaaacac 1740
cttcaagtgg acgggctaca tgtgctttcc tatagacctc atctggacctt ggaaggatta 1800
acccattca tgcacgagga ctcacgcca cggaaactggc gaaactacga ctagtaaaaa 1860
gactactgga cgaccgcctc ccacttagcg aggctcggct gtgtcgtggt agagtctccc 1920
tgtccggccc ttcattcaca ctgaggcgtt agaccgttct aataatcctc aaccctctca 1980
atgttcagac ctctctctg acccaactta aaactagacg agtttgaacg tccgctacat 2040
cttagtttag gacctgggag ggcctgttcc aggtatcgag agtgcaaaga gcgtcaacct 2100
cctcaagacg agaaggagag gcacttgcaac gtgagactgt gaccacacag gtatctgtag 2160
tcggccgttc tcgactctac accttcacct cacaagtatg 2200

```

&lt;210&gt; SEQ ID NO 67

&lt;211&gt; LENGTH: 868

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 67

```

Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Asn Met
1 5 10 15
Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Arg
20 25 30
Ala Met Leu Ser Leu Ile Asp Gly Lys Gly Pro Ile Arg Phe Val Leu
35 40 45
Ala Leu Leu Ala Phe Phe Arg Phe Thr Ala Ile Ala Pro Thr Arg Ala
50 55 60
Val Leu Asp Arg Trp Arg Gly Val Asn Lys Gln Thr Ala Met Lys His
65 70 75 80
Leu Leu Ser Phe Lys Lys Glu Leu Gly Thr Leu Thr Ser Ala Ile Asn
85 90 95
Arg Arg Ser Ser Lys Gln Lys Lys Arg Gly Gly Lys Thr Gly Ile Ala
100 105 110
Val Met Ile Gly Leu Ile Ala Ser Val Gly Ala Val Thr Leu Ser Asn
115 120 125
Phe Gln Gly Lys Val Met Met Thr Val Asn Ala Thr Asp Val Thr Asp
130 135 140
Val Ile Thr Ile Pro Thr Ala Ala Gly Lys Asn Leu Cys Ile Val Arg
145 150 155 160
Ala Met Asp Val Gly Tyr Met Cys Asp Asp Thr Ile Thr Tyr Glu Cys
165 170 175
Pro Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys

```

-continued

| 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Lys | Ser | Ala | Val | Tyr | Val | Arg | Tyr | Gly | Arg | Cys | Thr | Lys | Thr | Arg |
|     | 195 |     |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |
| His | Ser | Arg | Arg | Ser | Arg | Arg | Ser | Leu | Thr | Val | Gln | Thr | His | Gly | Glu |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |
| Ser | Thr | Leu | Ala | Asn | Lys | Lys | Gly | Ala | Trp | Met | Asp | Ser | Thr | Lys | Ala |
|     | 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     | 240 |
| Thr | Arg | Tyr | Leu | Val | Lys | Thr | Glu | Ser | Trp | Ile | Leu | Arg | Asn | Pro | Gly |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |
| Tyr | Ala | Leu | Val | Ala | Ala | Val | Ile | Gly | Trp | Met | Leu | Gly | Ser | Asn | Thr |
|     |     |     | 260 |     |     |     |     |     | 265 |     |     |     |     | 270 |     |
| Met | Gln | Arg | Val | Val | Phe | Val | Val | Leu | Leu | Leu | Leu | Val | Ala | Pro | Ala |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     |     | 285 |     |     |
| Tyr | Ser | Phe | Asn | Cys | Leu | Gly | Met | Ser | Asn | Arg | Asp | Phe | Leu | Glu | Gly |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |
| Val | Ser | Gly | Ala | Thr | Trp | Val | Asp | Leu | Val | Leu | Glu | Gly | Asp | Ser | Cys |
|     | 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     | 320 |
| Val | Thr | Ile | Met | Ser | Lys | Asp | Lys | Pro | Thr | Ile | Asp | Val | Lys | Met | Met |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |
| Asn | Met | Glu | Ala | Ala | Asn | Leu | Ala | Glu | Val | Arg | Ser | Tyr | Cys | Tyr | Leu |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     |     | 350 |     |
| Ala | Thr | Val | Ser | Asp | Leu | Ser | Thr | Lys | Ala | Ala | Cys | Pro | Ala | Met | Gly |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
| Glu | Ala | His | Asn | Asp | Lys | Arg | Ala | Asp | Pro | Ala | Phe | Val | Cys | Arg | Gln |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Gly | Val | Val | Asp | Arg | Gly | Trp | Gly | Asn | Gly | Cys | Gly | Leu | Phe | Gly | Lys |
|     | 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     | 400 |
| Gly | Ser | Ile | Asp | Thr | Cys | Ala | Lys | Phe | Ala | Cys | Ser | Thr | Lys | Ala | Ile |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Gly | Arg | Thr | Ile | Leu | Lys | Glu | Asn | Ile | Lys | Tyr | Glu | Val | Ala | Ile | Phe |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     |     | 430 |     |
| Val | His | Gly | Pro | Thr | Thr | Val | Glu | Ser | His | Gly | Asn | Tyr | Ser | Thr | Gln |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |
| Val | Gly | Ala | Thr | Gln | Ala | Gly | Arg | Phe | Ser | Ile | Thr | Pro | Ala | Ala | Pro |
|     | 450 |     |     |     |     | 455 |     |     |     |     |     | 460 |     |     |     |
| Ser | Tyr | Thr | Leu | Lys | Leu | Gly | Glu | Tyr | Gly | Glu | Val | Thr | Val | Asp | Cys |
|     | 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     | 480 |
| Glu | Pro | Arg | Ser | Gly | Ile | Asp | Thr | Asn | Ala | Tyr | Tyr | Val | Met | Thr | Val |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |
| Gly | Thr | Lys | Thr | Phe | Leu | Val | His | Arg | Glu | Trp | Phe | Met | Asp | Leu | Asn |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |
| Leu | Pro | Trp | Ser | Ser | Ala | Gly | Ser | Thr | Val | Trp | Arg | Asn | Arg | Glu | Thr |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Leu | Met | Glu | Phe | Glu | Glu | Pro | His | Ala | Thr | Lys | Gln | Ser | Val | Ile | Ala |
|     | 530 |     |     |     |     | 535 |     |     |     |     |     | 540 |     |     |     |
| Leu | Gly | Ser | Gln | Glu | Gly | Ala | Leu | His | Gln | Ala | Leu | Ala | Gly | Ala | Ile |
|     | 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     | 560 |
| Pro | Val | Glu | Phe | Ser | Ser | Asn | Thr | Val | Lys | Leu | Thr | Ser | Gly | His | Leu |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Lys | Cys | Arg | Val | Lys | Met | Glu | Lys | Leu | Gln | Leu | Lys | Gly | Thr | Thr | Tyr |
|     |     |     | 580 |     |     |     |     |     | 585 |     |     |     |     | 590 |     |
| Gly | Val | Cys | Ser | Lys | Ala | Phe | Lys | Phe | Leu | Gly | Thr | Pro | Ala | Asp | Thr |
|     |     | 595 |     |     |     |     |     | 600 |     |     |     |     |     | 605 |     |

-continued

Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly  
 610 615 620  
 Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr  
 625 630 635 640  
 Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr  
 645 650 655  
 Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser  
 660 665 670  
 Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His  
 675 680 685  
 Lys Ser Gly Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly  
 690 695 700  
 Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser  
 705 710 715 720  
 Val Gly Gly Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe  
 725 730 735  
 Gly Gly Ala Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln  
 740 745 750  
 Gly Leu Leu Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp  
 755 760 765  
 Arg Ser Ile Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe  
 770 775 780  
 Leu Ser Val Asn Val His Ala Asp Thr Gly Ile His Arg Gly Pro Ala  
 785 790 795 800  
 Thr Arg Thr Thr Thr Glu Ser Gly Lys Leu Ile Thr Asp Trp Cys Cys  
 805 810 815  
 Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Gln Thr Asp Ser Gly Cys  
 820 825 830  
 Trp Tyr Gly Met Glu Ile Arg Pro Gln Arg His Asp Glu Lys Thr Leu  
 835 840 845  
 Val Gln Ser Gln Val Asn Ala Tyr Asn Ala Asp Met Ile Asp Pro Phe  
 850 855 860  
 Gln Leu Gly Leu  
 865

&lt;210&gt; SEQ ID NO 68

&lt;211&gt; LENGTH: 2700

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 68

```

agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta 60
acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc 120
ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtggt gtccttgatt 180
ggacttaaga gggctatggt gagcctgac gacggcaagg ggccaatacg atttgtgttg 240
gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga 300
tggagagggtg tgaacaaaca aacagcgatg aaacaccttc tgagtttcaa gaaggaacta 360
gggaccttga ccagtgtat caatcggcgg agctcaaac aaaagaaaag aggaggaaaag 420
accggaattg cagtcatgat tggcctgac gccagcgtag gacagttac cctctctaac 480
ttccaagggg agtgtatgat gacggtaat gctactgacg tcacagatgt catcacgatt 540

```

-continued

---

```

ccaacagctg ctggaaaagaa cctatgcatt gtcagagcaa tggatgtggg atacatgtgc 600
gatgatacta tcaactatga atgcccagtg ctgtcggctg gtaatgatcc agaagacatc 660
gactgttggg gcacaaaagtc agcagtcctac gtcaggatg gaagatgcac caagacacgc 720
cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaaag cactctagcg 780
aacaagaagg gggcttggat ggacagcacc aaggccacaa ggtatttggg aaaaacagaa 840
tcatggatct tgaggaaccc tggatatgcc ctggtggcag ccgtcattgg ttggatgctt 900
gggagcaaca ccatgcagag agttgtgtt gtctgtctat tgcttttggg ggccccagct 960
tacagcttta actgccttgg aatgagcaac agagacttct tggaaaggag gtctggagca 1020
acatgggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc taaggacaag 1080
cctaccatcg atgtgaagat gatgaatag gaggcggcca acctggcaga ggtccgcagt 1140
tattgctatt tggctacogt cagcgatctc tccaccaaag ctgctgccc ggccatggga 1200
gaagctcaca atgacaaaag tgctgaccca gcttttgtgt gcagacaagg agtgggtggac 1260
aggggctggg gcaacggctg cggactattt ggcaaaggaa gcattgacac atgcccacaa 1320
tttgctgct ctaccaaggc aataggaaga accattttga aagagaatat caagtacgaa 1380
gtggccattt ttgtccatgg accaactact gtggagtgcg acggaaacta ctccacacag 1440
gttgagacca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta 1500
aagcttgag aatatggaga ggtgacagtg gactgtgaac cacggtcagg gattgacacc 1560
aatgcatact acgtgatgac tgttggaaaca aagacgttct tggtcocacg tgagtggttc 1620
atggacctca acctcccttg gagcagtgtc ggaagtactg tgtggaggaa cagagagacg 1680
ttaatggagt ttgaggaacc acacgccacg aagcagtctg tgatagcatt gggctcacia 1740
gagggagctc tgcacaaagc tttggctgga gccattcctg tggaaatttc aagcaacact 1800
gtcaagttga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag 1860
ggaacaacct atggcgtctg ttcaaaggct ttcaagtctc ttgggactcc cgcagacaca 1920
ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt 1980
cctatctcgt cagtggcttc attgaaagc ctaacgccag tgggcagatt ggtcactgtc 2040
aaccctttg tttcagtggc cacggccaac gctaaggctc tgattgaatt ggaaccaccc 2100
tttgagact catacatagt ggtgggcaga ggagaacaac agatcaatca ccactggcac 2160
aagtctggaa gcagcattgg caaagcctt acaaccaccc tcaaaggagc gcagagacta 2220
gccgctctag gagacacagc ttgggacttt ggatcagttg gaggggtgtt cacctcagtt 2280
gggaaggctg tccatcaagt gttcggagga gcattccgct cactgttcgg aggcattgct 2340
tggataacgc aaggattgct gggggctctc ctgttggtga tgggcatcaa tgetcgtgac 2400
aggtccatag ctctcacggt tctcgcagtt ggaggagttc tgctcttctc ctccgtgaac 2460
gtgcacgctg aactgggat ccaccgtgga cctgccactc gcaccaccac agagagcgga 2520
aagttgataa cagattgggt ctgcaggagc tgcaccttac caccactgcg ctaccaaact 2580
gacagcggct gttggtatgg tatggagatc agaccacaga gacatgatga aaagacctc 2640
gtgcagtcac aagtgaatgc ttataatgct gatatgattg accttttca gttgggcctt 2700

```

&lt;210&gt; SEQ ID NO 69

&lt;211&gt; LENGTH: 2700

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

-continued

---

<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 69

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tcatcaagcg gacacactcg actgtttgaa tcatcacaaa cactoctaat tgttgtaat   | 60   |
| tgtgtcacgc tcgacaaaga atcgtgcttc tagagctaca gattctttgg tcctcccggg  | 120  |
| ccgttctcgg cccgacagtt atacgatttt gcgccttacg gggcgcaaaa caggaactaa  | 180  |
| cctgaattct cccgatacaa ctggactag ctgccttcc cgggttatgc taaacacaac    | 240  |
| cgagagaacc gcaagaagtc caagtgtcgt taacgaggct gggctcgtca cgacctagct  | 300  |
| acctctccac acttgtttgt ttgtcgtac tttgtggaag actcaaagtt ctctcttgat   | 360  |
| ccctggaact ggtaacgata gtttagccgc tcgagttttg tttctcttcc tcctccttc   | 420  |
| tggccttaac gtcagtaact accggactag cggtcgcctc ctctcaatg ggagagattg   | 480  |
| aaggttccct tccactacta ctgccattta cgatgactgc agtgtctaca gtagtgctaa  | 540  |
| ggttgtcgac gacctttctt ggatacgtaa cagtctcgtt acctacacc tatgtacacg   | 600  |
| ctactatgat agtgaatact tacgggtcac gacagccgac cttactagg tctctgtag    | 660  |
| ctgacaacca cgtgtttcag tcgtcagatg cagtcctaac ctctacgtg gttctgtgcg   | 720  |
| gtgagttctg cgtcagcctc cagtgactgt cacgtctgtg tgctcttcc gtgagatcgc   | 780  |
| ttgttcttcc cccgaaccta cctgtcgtgg ttccgggttt ccataaacca tttttgtctt  | 840  |
| agtacctaga actccttggg acctatacgg gaccaccgtc ggcagtaacc aaacctacgaa | 900  |
| ccctcgttgt ggtacgtctc tcaacacaaa cagcagata acgaaaacca cgggggtcga   | 960  |
| atgtcgaaat tgacggaacc ttactcgttg tctctgaaga acctctca cagacctcgt    | 1020 |
| tgtaccacc taaaccaaga gcttccgctg tcgacgcact gatagtacag attcctgttc   | 1080 |
| ggatggtagc tacacttota ctacttatac ctccgccggt tggaccgtct ccaggcgtca  | 1140 |
| ataacgataa accgatggca gtcgctagag aggtggtttc gacgcacggg ccggtaccct  | 1200 |
| cttcgagtgt tactgtttgc acgactgggt cgaaaacaca cgtctgttcc tcaccacctg  | 1260 |
| tccccgacc cgttgccgac gctgataaa ccgtttcctt cgtaactgtg tacgaggttt    | 1320 |
| aaacggacga gatggttccg ttatccttct tggtaaaact ttctcttata gttcatgctt  | 1380 |
| caccggtaaa aacaggtaacc tggttgatga cacctcagcg tgctttgat gaggtgtgtc  | 1440 |
| caacctcggg gagtccgtcc ctctaagtcg tagtgaggac gccgcggaag tatgtgtgat  | 1500 |
| ttcgaacctc ttatacctct cactgtcac ctgacacttg gtgcccagtc ctaactgtgg   | 1560 |
| ttacgtatga tgcactactg acaaccttgt ttctgcaaga accaggtagc actcaccaag  | 1620 |
| tacctggagt tggagggaac ctgcgcagca ccttcatgac acacctcctt gtctctctgc  | 1680 |
| aattacctca aactccttg tgtgcgggtc ttcgtcagac actatcgtaa cccgagtgtt   | 1740 |
| ctccctcgag acgtagtctg aaaccgacct cggtaaggac accttaaaag ttcgtgtgta  | 1800 |
| cagttcaact gcagcccagt aaacttcaca tctcacttct acctttttaa cgtcaacttc  | 1860 |
| ccttggtgga taccgcagac aagtttccga aagttcaaag aacctgagg gcgtctgtgt   | 1920 |
| ccagtgccgt gacaccacaa ccttaacgtc atgtgaccgt gcctacctgg aacgtttcaa  | 1980 |
| ggatagagca gtcaccgaag taacttctgt gattgctgtc acccgtctaa ccagtgcag   | 2040 |
| ttgggaaaac aaagtcaccg gtgcccgttg cgattccagg actaacttaa ccttggtggg  | 2100 |
| aaacctctga gtatgtatca ccacctctc cctctgttg tctagttagt ggtgaccgtg    | 2160 |
| ttcagacctt cgtcgtaac gtttcggaaa tgttggtggg agtttctctg cgtctctgat   | 2220 |
| cggcgagatc ctctgtgtcg aacctgaaa cctagtcaac ctccccacaa gtggagtcaa   | 2280 |

-continued

---

```

cccttccgac aggtagtcca caagcctcct cgtaaggcga gtgacaagcc tccgtacagg 2340
acctattgog ttctaacga cccccgagag gacaacacct acccgtagtt acgagcactg 2400
tccaggtatc gagagtgcaa agagcgtcaa cctcctcaag acgagaagga gaggcacttg 2460
cacgtgcgac tgtgacccta ggtggcacct ggacggtgag cgtgggtggtg tctctcgct 2520
ttcaactatt gtctaaccac gacgtcctcg acgtggaatg gtggtgacgc gatggtttga 2580
ctgtcgccga caaccatacc atacctctag tctgggtgtct ctgtactact tttctgggag 2640
cacgtcagtg ttcacttacg aatattacga ctatactaac tgggaaaagt caaccgggaa 2700

```

```

<210> SEQ ID NO 70
<211> LENGTH: 734
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 70

```

```

Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Tyr Leu
 1 5 10 15
Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Arg
 20 25 30
Ala Met Leu Ser Leu Ile Asp Gly Lys Gly Pro Ile Arg Phe Val Leu
 35 40 45
Ala Leu Leu Ala Phe Phe Arg Phe Thr Ala Ile Ala Pro Thr Arg Ala
 50 55 60
Val Leu Asp Arg Trp Arg Gly Val Asn Lys Gln Thr Ala Met Lys His
 65 70 75 80
Leu Leu Ser Phe Lys Lys Glu Leu Gly Thr Leu Thr Ser Ala Ile Asn
 85 90 95
Arg Arg Ser Ser Lys Gln Lys Lys Arg Gly Gly Glu Leu Leu Ile Leu
100 105 110
Lys Ala Asn Ala Ile Thr Thr Ile Leu Thr Ala Val Thr Phe Cys Phe
115 120 125
Ala Ser Gly Gln Asn Ile Thr Glu Glu Phe Tyr Gln Ser Thr Cys Ser
130 135 140
Ala Val Ser Lys Gly Tyr Leu Ser Ala Leu Arg Thr Gly Trp Tyr Thr
145 150 155 160
Ser Val Ile Thr Ile Glu Leu Ser Asn Ile Lys Glu Asn Lys Cys Asn
165 170 175
Gly Thr Asp Ala Lys Val Lys Leu Ile Lys Gln Glu Leu Asp Lys Tyr
180 185 190
Lys Asn Ala Val Thr Glu Leu Gln Leu Leu Met Gln Ser Thr Pro Pro
195 200 205
Thr Asn Asn Arg Ala Arg Arg Glu Leu Pro Arg Phe Met Asn Tyr Thr
210 215 220
Leu Asn Asn Ala Lys Lys Thr Asn Val Thr Leu Ser Lys Lys Arg Lys
225 230 235 240
Arg Arg Phe Leu Gly Phe Leu Leu Gly Val Gly Ser Ala Ile Ala Ser
245 250 255
Gly Val Ala Val Ser Lys Val Leu His Leu Glu Gly Glu Val Asn Lys
260 265 270
Ile Lys Ser Ala Leu Leu Ser Thr Asn Lys Ala Val Val Ser Leu Ser
275 280 285

```

-continued

---

Asn Gly Val Ser Val Leu Thr Ser Lys Val Leu Asp Leu Lys Asn Tyr  
 290 295 300  
 Ile Asp Lys Gln Leu Leu Pro Ile Val Asn Lys Gln Ser Cys Ser Ile  
 305 310 315 320  
 Ser Asn Ile Glu Thr Val Ile Glu Phe Gln Gln Lys Asn Asn Arg Leu  
 325 330 335  
 Leu Glu Ile Thr Arg Glu Phe Ser Val Asn Ala Gly Val Thr Thr Pro  
 340 345 350  
 Val Ser Thr Tyr Met Leu Thr Asn Ser Glu Leu Leu Ser Leu Ile Asn  
 355 360 365  
 Asp Met Pro Ile Thr Asn Asp Gln Lys Lys Leu Met Ser Asn Asn Val  
 370 375 380  
 Gln Ile Val Arg Gln Gln Ser Tyr Ser Ile Met Ser Ile Ile Lys Glu  
 385 390 395 400  
 Glu Val Leu Ala Tyr Val Val Gln Leu Pro Leu Tyr Gly Val Ile Asp  
 405 410 415  
 Thr Pro Cys Trp Lys Leu His Thr Ser Pro Leu Cys Thr Thr Asn Thr  
 420 425 430  
 Lys Glu Gly Ser Asn Ile Cys Leu Thr Arg Thr Asp Arg Gly Trp Tyr  
 435 440 445  
 Cys Asp Asn Ala Gly Ser Val Ser Phe Phe Pro Gln Ala Glu Thr Cys  
 450 455 460  
 Lys Val Gln Ser Asn Arg Val Phe Cys Asp Thr Met Asn Ser Leu Thr  
 465 470 475 480  
 Leu Pro Ser Glu Ile Asn Leu Cys Asn Val Asp Ile Phe Asn Pro Lys  
 485 490 495  
 Tyr Asp Cys Lys Ile Met Thr Ser Lys Thr Asp Val Ser Ser Ser Val  
 500 505 510  
 Ile Thr Ser Leu Gly Ala Ile Val Ser Cys Tyr Gly Lys Thr Lys Cys  
 515 520 525  
 Thr Ala Ser Asn Lys Asn Arg Gly Ile Ile Lys Thr Phe Ser Asn Gly  
 530 535 540  
 Cys Asp Tyr Val Ser Asn Lys Gly Met Asp Thr Val Ser Val Gly Asn  
 545 550 555 560  
 Thr Leu Tyr Tyr Val Asn Lys Gln Glu Gly Lys Ser Leu Tyr Val Lys  
 565 570 575  
 Gly Glu Pro Ile Ile Asn Phe Tyr Asp Pro Leu Val Phe Pro Ser Asp  
 580 585 590  
 Glu Phe Asp Ala Ser Ile Ser Gln Val Asn Glu Lys Ile Asn Gln Ser  
 595 600 605  
 Leu Ala Phe Ile Arg Lys Ser Asp Glu Leu Leu His Asn Val Asn Ala  
 610 615 620  
 Gly Lys Ser Thr Thr Asn Ile Met Ile Thr Thr Ile Ile Ile Val Ile  
 625 630 635 640  
 Ile Val Ile Leu Leu Ser Leu Ile Ala Val Gly Leu Leu Leu Tyr Cys  
 645 650 655  
 Lys Ala Arg Ser Thr Pro Val Thr Leu Ser Lys Asp Gln Leu Ser Gly  
 660 665 670  
 Ile Asn Asn Ile Ala Phe Ser Asn Asn Phe Asp Leu Leu Lys Leu Ala  
 675 680 685  
 Gly Asp Val Glu Ser Asn Pro Gly Pro Ala Arg Asp Arg Ser Ile Ala  
 690 695 700  
 Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser Val Asn

-continued

| 705                                                                                                                                                                                            | 710        | 715        | 720        |                            |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|------------|----------------------------|
| Val His Ala Asp Thr Gly Cys Ala Ile Asp Ile Ser Arg Gln                                                                                                                                        |            |            |            |                            |
|                                                                                                                                                                                                | 725        | 730        |            |                            |
| <210> SEQ ID NO 71<br><211> LENGTH: 2298<br><212> TYPE: DNA<br><213> ORGANISM: Artificial Sequence<br><220> FEATURE:<br><223> OTHER INFORMATION: Synthetic Construct<br><br><400> SEQUENCE: 71 |            |            |            |                            |
| agtagttcgc                                                                                                                                                                                     | ctgtgtgagc | tgacaaactt | agtagtgttt | gtgaggatta acaacaatta 60   |
| acacagtgcg                                                                                                                                                                                     | agctgtttct | tagcacgaag | atctcgatgt | ctaagaaacc aggagggccc 120  |
| ggcaagagcc                                                                                                                                                                                     | gggctgtota | tttgctaaaa | cgcggaatgc | cccgcgtgtt gtccttgatt 180  |
| ggacttaaga                                                                                                                                                                                     | gggctatggt | gagcctgatc | gacggcaagg | ggccaatacg atttgtgttg 240  |
| gctctcttgg                                                                                                                                                                                     | cgttcttcag | gttcacagca | attgctccga | cccgagcagt gctggatcga 300  |
| tggagaggtg                                                                                                                                                                                     | tgaacaaaca | aacagcgtat | aaacaccttc | tgagtttcaa gaaggaacta 360  |
| gggaccttga                                                                                                                                                                                     | ccagtgtcat | caatcggcgg | agctcaaagc | aaaagaagcg agggggcgag 420  |
| ttgctaatac                                                                                                                                                                                     | tcaaaagaaa | tgcaattacc | acaatcctca | ctgcagtcac attttgtttt 480  |
| gcttctgtgc                                                                                                                                                                                     | aaaacatcac | tgaagaattt | tatcaatcaa | catgcagtgc agttagcaaa 540  |
| ggctatctta                                                                                                                                                                                     | gtgctctgag | aactggttgg | tataccagtg | ttataactat agaattaagt 600  |
| aatatcaagg                                                                                                                                                                                     | aaaataagtg | taatggaaca | gatgctaagg | taaaattgat aaaacaagaa 660  |
| ttagataaat                                                                                                                                                                                     | ataaaaatgc | tgtaacagaa | ttgcagttgc | tcatgcaaag cacaccacca 720  |
| acaaacaatc                                                                                                                                                                                     | gagccagaag | agaactacca | aggtttatga | attatacact caacaatgcc 780  |
| aaaaaaaaacca                                                                                                                                                                                   | atgtaacatt | aagcaagaaa | aggaaaagaa | gatttcttgg ttttttgta 840   |
| gggtgttgat                                                                                                                                                                                     | ctgcaatcgc | cagtggcggt | gctgtatcta | aggctctgca cctagaaggg 900  |
| gaagtgaaca                                                                                                                                                                                     | agatcaaaag | tgctctacta | tccacaaaca | aggctgtagt cagottatca 960  |
| aatggagtta                                                                                                                                                                                     | gtgtcttaac | cagcaaagtg | ttagacctca | aaaactatat agataaaca 1020  |
| ttgttaccta                                                                                                                                                                                     | ttgtgaacaa | gcaaagctgc | agcatatcaa | atatagaaac tgtgatagag 1080 |
| ttccaacaaa                                                                                                                                                                                     | agaacaacag | actactagag | attaccaggg | aatttagtgt taatgcaggt 1140 |
| gtaactacac                                                                                                                                                                                     | ctgtaagcac | ttacatgtta | actaatagtg | aattattgtc attaataaat 1200 |
| gatatgccta                                                                                                                                                                                     | taacaaatga | tcagaaaaag | ttaatgtcca | acaatgttca aatagttaga 1260 |
| cagcaaagtt                                                                                                                                                                                     | actctatcat | gtccataata | aaagaggaag | tcttagcata tgtagtacaa 1320 |
| ttaccactat                                                                                                                                                                                     | atgggtttat | agatacacc  | tggttgaaac | tacacacatc ccctctatgt 1380 |
| acaaccaaca                                                                                                                                                                                     | caaaagaagg | gtccaacatc | tgtttaacaa | gaactgacag aggatggtag 1440 |
| tgtgacaatg                                                                                                                                                                                     | caggatcagt | atctttcttc | ccacaagctg | aaacatgtaa agttcaatca 1500 |
| aatcgagtat                                                                                                                                                                                     | tttgtgacac | aatgaacagt | ttaacattac | caagtgaaat aaatctctgc 1560 |
| aatgttgaca                                                                                                                                                                                     | tattcaaccc | caaatatgat | tgtaaaatta | tgacttcaaa aacagatgta 1620 |
| agcagctcog                                                                                                                                                                                     | ttatcacatc | tctaggagcc | attgtgtcat | gctatggcaa aactaaatgt 1680 |
| acagcatcca                                                                                                                                                                                     | ataaaaatcg | tggaatcata | aagacatttt | ctaacgggtg cgattatgta 1740 |
| tcaataaag                                                                                                                                                                                      | ggatggacac | tgtgtctgta | ggtaacacat | tatattatgt aaataagcaa 1800 |
| gaaggtaaaa                                                                                                                                                                                     | gtctctatgt | aaaaggtgaa | ccaataataa | atctctatga cccattagta 1860 |
| ttcccctctg                                                                                                                                                                                     | atgaatttga | tgcatcaata | tctcaagtca | acgagaagat taaccagagc 1920 |

-continued

---

|                                                                   |      |
|-------------------------------------------------------------------|------|
| ctagcattta ttcgtaaate cgatgaatta ttacataatg taaatgctgg taaatccacc | 1980 |
| acaaatatca tgataactac tataattata gtgattatag taatattgtt atcattaatt | 2040 |
| gctgttgac tgctcttata ctgtaaggcc agaagcacac cagtcacact aagcaaagat  | 2100 |
| caactgagtg gtataaataa tattgcattt agtaacaatt ttgatctgct caaacttgca | 2160 |
| ggcgatgtag aatcaaatec tggaccogcc cgggacaggt ccatagctct cactgttctc | 2220 |
| gcagttggag gagttctgct ctctctctcc gtgaacgtgc acgctgacac tgggtgtgcc | 2280 |
| atagacatca gccggcaa                                               | 2298 |

&lt;210&gt; SEQ ID NO 72

&lt;211&gt; LENGTH: 2298

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 72

|                                                                   |      |
|-------------------------------------------------------------------|------|
| tcatcaagcg gacacactcg actgttgaa tcatcacaaa cactcctaat tgttgtaaat  | 60   |
| tgtgtcacgc tcgacaaaga atcgtgcttc tagagctaca gattctttgg tcctcccggg | 120  |
| ccgttctcgg cccgacagat aaacgatttt gcgccttacg gggcgcacaa caggaactaa | 180  |
| cctgaattct cccgatacaa ctccgactag ctgccgttcc ccggttatgc taaacacaac | 240  |
| cgagagaacc gcaagaagtc caagtgtcgt taacgaggct gggctcgtca cgacctagct | 300  |
| acctctccac acttgtttgt ttgtcgctac tttgtggaag actcaaagtt ctctcttgat | 360  |
| ccctggaact ggtcacgata gttagccgcc tcgagtttcc tttctctgc tccccgctc   | 420  |
| aacgattagg agttctgttt acgttaatgg tgttaggagt gacgtcagtg taaaacaaaa | 480  |
| cgaagaccag tttttagtag actctctaaa atagttagtt gtacgtcacg tcaatcgttt | 540  |
| ccgatagaat caccgactc ttgaccaacc atatggtcac aatattgata tcttaattca  | 600  |
| ttatagttcc ttttatccac attacctgt ctacgattcc attttaacta tttgttctt   | 660  |
| aatctattta tatttttacg acattgtctt aacgtcaacg agtacgttcc gtgtggtggt | 720  |
| tgtttgtag ctccgctctc tcttgatggt tccaaact taatatgtga gttgttacgg    | 780  |
| ttttttggt tacattgtaa ttcgttcttt tcttttctt ctaaagaacc aaaaaacaat   | 840  |
| ccacaacctc gacgttagcg gtcaccgcaa cgacatagat tccaggacgt ggatcttccc | 900  |
| cttcacttgt tctagtttcc acgagatgat aggtgtttgt tccgacatca gtcgaatagt | 960  |
| ttacctcaat cacagaattg gtcgtttcac aatctggagt ttttgatata tctattgtt  | 1020 |
| aacaatggat aacactgtt cgtttcgacg tcgtatagtt tatacttttg acactatctc  | 1080 |
| aaggttggtt tcttggtgtc tgatgatctc taatggccc ttaaatcaca attacgtcca  | 1140 |
| cattgatgty gacattcgtg aatgtacaat tgattatcac ttaataacag taattagtta | 1200 |
| ctatacggat attgttact agtcttttcc aattacaggt tgttacaagt ttatcaatct  | 1260 |
| gtcgtttcaa tgagatagta caggattat tttctcttc agaatcgtat acatcatggt   | 1320 |
| aatggtgata taccacaata tctatgtggg acaacctttg atgtgtgtag gggagataca | 1380 |
| tgttggttgt gtttctctcc caggttgtag acaaatgtt cttgactgtc tcctaccatg  | 1440 |
| acactgttac gtccatgca tagaaagaag ggtgttcgac tttgtacatt tcaagttagt  | 1500 |
| ttagctcata aaacactgty ttactgttca aattgtaatg gttcacttta tttagagacg | 1560 |
| ttacaactgt ataagttggg gtttatacta acattttaat actgaagttt ttgtctacat | 1620 |
| tcgtcagggc aatagttag agatcctcgg taacacagta cgataccgtt ttgatttaca  | 1680 |

-continued

---

```

tgctgtaggt tatttttagc accttagtat ttctgtaaaa gattgcccac gctaatacat 1740
agtttatttc cctacctgtg acacagacat ccattgtgta atataatata tttattcggt 1800
cttccatttt cagagatata ttttccactt ggttattatt taaagatact gggtaatcat 1860
aaggggagac tacttaaact acgtagttat agagtccagt tgetcttcta attggctcgc 1920
gatcgtaaat aagcatttag gctacttaat aatgtattac atttacgacc atttaggtgg 1980
tgtttatagt actattgatg atattaatat cactaataatc attataacaa tagtaattaa 2040
cgacaacctg acgagaatat gacattccgg tcttcgtgtg gtcagtgtga ttcgtttcta 2100
gttgactcac catatttatt ataacgtaaa tcattgtaa aactagacga gtttgaacgt 2160
ccgctacatc ttagttagg acctgggccc gccctgtcca ggtatcgaga gtgcaaagag 2220
cgtcaacctc ctcaagacga gaaggagagg cacttgcaag tgcgactgtg acccacacgg 2280
tatctgtagt cggccggt 2298

```

```

<210> SEQ ID NO 73
<211> LENGTH: 667
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 73

```

```

Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Asn Met
1 5 10 15
Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Gln
20 25 30
Lys Lys Arg Gly Gly Glu Leu Leu Ile Leu Lys Ala Asn Ala Ile Thr
35 40 45
Thr Ile Leu Thr Ala Val Thr Phe Cys Phe Ala Ser Gly Gln Asn Ile
50 55 60
Thr Glu Glu Phe Tyr Gln Ser Thr Cys Ser Ala Val Ser Lys Gly Tyr
65 70 75 80
Leu Ser Ala Leu Arg Thr Gly Trp Tyr Thr Ser Val Ile Thr Ile Glu
85 90 95
Leu Ser Asn Ile Lys Glu Asn Lys Cys Asn Gly Thr Asp Ala Lys Val
100 105 110
Lys Leu Ile Lys Gln Glu Leu Asp Lys Tyr Lys Asn Ala Val Thr Glu
115 120 125
Leu Gln Leu Leu Met Gln Ser Thr Pro Pro Thr Asn Asn Arg Ala Arg
130 135 140
Arg Glu Leu Pro Arg Phe Met Asn Tyr Thr Leu Asn Asn Ala Lys Lys
145 150 155 160
Thr Asn Val Thr Leu Ser Lys Lys Arg Lys Arg Arg Phe Leu Gly Phe
165 170 175
Leu Leu Gly Val Gly Ser Ala Ile Ala Ser Gly Val Ala Val Ser Lys
180 185 190
Val Leu His Leu Glu Gly Glu Val Asn Lys Ile Lys Ser Ala Leu Leu
195 200 205
Ser Thr Asn Lys Ala Val Val Ser Leu Ser Asn Gly Val Ser Val Leu
210 215 220
Thr Ser Lys Val Leu Asp Leu Lys Asn Tyr Ile Asp Lys Gln Leu Leu
225 230 235 240
Pro Ile Val Asn Lys Gln Ser Cys Ser Ile Ser Asn Ile Glu Thr Val

```

-continued

| 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Glu | Phe | Gln | Gln | Lys | Asn | Asn | Arg | Leu | Leu | Glu | Ile | Thr | Arg | Glu |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     |     | 270 |     |
| Phe | Ser | Val | Asn | Ala | Gly | Val | Thr | Thr | Pro | Val | Ser | Thr | Tyr | Met | Leu |
|     |     | 275 |     |     |     |     |     | 280 |     |     |     |     | 285 |     |     |
| Thr | Asn | Ser | Glu | Leu | Leu | Ser | Leu | Ile | Asn | Asp | Met | Pro | Ile | Thr | Asn |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |
| Asp | Gln | Lys | Lys | Leu | Met | Ser | Asn | Asn | Val | Gln | Ile | Val | Arg | Gln | Gln |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |
| Ser | Tyr | Ser | Ile | Met | Ser | Ile | Ile | Lys | Glu | Glu | Val | Leu | Ala | Tyr | Val |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     |     | 335 |
| Val | Gln | Leu | Pro | Leu | Tyr | Gly | Val | Ile | Asp | Thr | Pro | Cys | Trp | Lys | Leu |
|     |     |     | 340 |     |     |     |     |     | 345 |     |     |     |     | 350 |     |
| His | Thr | Ser | Pro | Leu | Cys | Thr | Thr | Asn | Thr | Lys | Glu | Gly | Ser | Asn | Ile |
|     |     | 355 |     |     |     |     |     | 360 |     |     |     |     | 365 |     |     |
| Cys | Leu | Thr | Arg | Thr | Asp | Arg | Gly | Trp | Tyr | Cys | Asp | Asn | Ala | Gly | Ser |
| 370 |     |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Val | Ser | Phe | Phe | Pro | Gln | Ala | Glu | Thr | Cys | Lys | Val | Gln | Ser | Asn | Arg |
| 385 |     |     |     |     | 390 |     |     |     |     |     | 395 |     |     |     | 400 |
| Val | Phe | Cys | Asp | Thr | Met | Asn | Ser | Leu | Thr | Leu | Pro | Ser | Glu | Ile | Asn |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     |     | 415 |
| Leu | Cys | Asn | Val | Asp | Ile | Phe | Asn | Pro | Lys | Tyr | Asp | Cys | Lys | Ile | Met |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     |     | 430 |     |
| Thr | Ser | Lys | Thr | Asp | Val | Ser | Ser | Ser | Val | Ile | Thr | Ser | Leu | Gly | Ala |
|     |     | 435 |     |     |     |     |     | 440 |     |     |     |     | 445 |     |     |
| Ile | Val | Ser | Cys | Tyr | Gly | Lys | Thr | Lys | Cys | Thr | Ala | Ser | Asn | Lys | Asn |
| 450 |     |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Arg | Gly | Ile | Ile | Lys | Thr | Phe | Ser | Asn | Gly | Cys | Asp | Tyr | Val | Ser | Asn |
| 465 |     |     |     |     | 470 |     |     |     |     |     | 475 |     |     |     | 480 |
| Lys | Gly | Met | Asp | Thr | Val | Ser | Val | Gly | Asn | Thr | Leu | Tyr | Tyr | Val | Asn |
|     |     |     | 485 |     |     |     |     |     | 490 |     |     |     |     |     | 495 |
| Lys | Gln | Glu | Gly | Lys | Ser | Leu | Tyr | Val | Lys | Gly | Glu | Pro | Ile | Ile | Asn |
|     |     |     | 500 |     |     |     |     |     | 505 |     |     |     |     | 510 |     |
| Phe | Tyr | Asp | Pro | Leu | Val | Phe | Pro | Ser | Asp | Glu | Phe | Asp | Ala | Ser | Ile |
|     |     | 515 |     |     |     |     |     | 520 |     |     |     |     | 525 |     |     |
| Ser | Gln | Val | Asn | Glu | Lys | Ile | Asn | Gln | Ser | Leu | Ala | Phe | Ile | Arg | Lys |
| 530 |     |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Ser | Asp | Glu | Leu | Leu | His | Asn | Val | Asn | Ala | Gly | Lys | Ser | Thr | Thr | Asn |
| 545 |     |     |     |     | 550 |     |     |     |     |     | 555 |     |     |     | 560 |
| Ile | Met | Ile | Thr | Thr | Ile | Ile | Ile | Val | Ile | Ile | Val | Ile | Leu | Leu | Ser |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     |     | 575 |
| Leu | Ile | Ala | Val | Gly | Leu | Leu | Leu | Tyr | Cys | Lys | Ala | Arg | Ser | Thr | Pro |
|     |     |     | 580 |     |     |     |     |     | 585 |     |     |     |     | 590 |     |
| Val | Thr | Leu | Ser | Lys | Asp | Gln | Leu | Ser | Gly | Ile | Asn | Asn | Ile | Ala | Phe |
|     |     | 595 |     |     |     |     |     | 600 |     |     |     |     | 605 |     |     |
| Ser | Asn | Asn | Phe | Asp | Leu | Leu | Lys | Leu | Ala | Gly | Asp | Val | Glu | Ser | Asn |
|     | 610 |     |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |
| Pro | Gly | Pro | Ala | Arg | Asp | Arg | Ser | Ile | Ala | Leu | Thr | Phe | Leu | Ala | Val |
| 625 |     |     |     |     | 630 |     |     |     |     |     | 635 |     |     |     | 640 |
| Gly | Gly | Val | Leu | Leu | Phe | Leu | Ser | Val | Asn | Val | His | Ala | Asp | Thr | Gly |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     |     | 655 |
| Cys | Ala | Ile | Asp | Ile | Ser | Arg | Gln | Glu | Leu | Arg |     |     |     |     |     |
|     |     |     | 660 |     |     |     |     |     | 665 |     |     |     |     |     |     |

-continued

---

```

<210> SEQ ID NO 74
<211> LENGTH: 2097
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 74
agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta 60
acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc 120
ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtggt gtccttgatt 180
ggacttaagc aaaagaagcg agggggcgag ttgctaatec tcaaagcaaa tgcaattacc 240
acaatcctca ctgcagtcac attttgtttt gcttctggtc aaaacatcac tgaagaattt 300
tatcaatcaa catgcagtgc agttagcaaa ggctatctta gtgctctgag aactggttgg 360
tataccagtg ttataactat agaattaagt aatatcaagg aaaataagtg taatggaaca 420
gatgctaagg taaaattgat aaaacaagaa ttagataaat ataaaaatgc tgtaacagaa 480
ttgcagttgc tcatgcaaag cacaccacca acaacaatc gagccagaag agaactacca 540
aggtttatga attatacact caacaatgcc aaaaaacca atgtaacatt aagcaagaaa 600
aggaaaagaa gatttcttgg ttttttgta ggtgttggat ctgcaatgc cagtggcgtt 660
gctgtatcta aggtcctgca cctagaaggg gaagtgaaca agatcaaaag tgcttacta 720
tccacaaaaca aggtctgtagt cagcctatca aatggagtta gtgtcttaac cagcaaagtg 780
ttagacctca aaaactatat agataaacia ttgttaccta ttgtgaacia gcaaagctgc 840
agcatatcaa atatagaaac tgtgatagag ttccaacaaa agaacaacag actactagag 900
attaccaggg aatttagtgt taatgcaggt gtaactacac ctgtaagcac ttacatgta 960
actaatagtg aattattgtc attaatcaat gatatgccta taacaaatga tcagaaaaag 1020
ttaatgtcca acaatgttca aatagttaga cagcaaagtt actctatcat gtccataata 1080
aaagaggaag tcttagcata tgtagtacaa ttaccactat atggtgttat agatacacc 1140
tgttggaaac tacacacatc cctctatgt acaaccaaca caaaagaagg gtccaacatc 1200
tgtttaacaa gaactgacag aggatggtag tgtgacaatg caggatcagt atctttcttc 1260
ccacaagctg aaacatgtaa agttcaatca aatcgagtat tttgtgacac aatgaacagt 1320
ttaacattac caagtgaat aaatctctgc aatgttgaca tattcaaccc caaatatgat 1380
tgtaaaatta tgacttcaaa aacagatgta agcagctccg ttatccatc tctaggagcc 1440
attgtgcat gctatggcaa aactaaatgt acagcatcca ataaaaatcg tggaatcata 1500
aagacatttt ctaacgggtg cgattatgta tcaataaag ggatggacac tgtgtctgta 1560
ggtaacacat tatattatgt aaataagcaa gaaggtaaaa gtctctatgt aaaagtgaa 1620
ccaataataa atttctatga cccattagta ttcccctctg atgaatttga tgcataata 1680
tctcaagtca acgagaagat taaccagagc ctagcattta ttcgtaaatc cgatgaatta 1740
ttacataatg taaatgctgg taaatccacc acaaatatca tgataactac tataattata 1800
gtgattatag taatattggt atcattaatt gctgttggac tgctcttata ctgtaaggcc 1860
agaagcacac cagtcacact aagcaaagat caactgagtg gtataaataa tattgcattt 1920
agtaacaatt ttgatctgct caaacttgca ggcatgtag aatcaaatcc tggaccgcc 1980
egggacaggt ccatagctct caggtttctc gcagttggag gagttctgct ctctctctcc 2040

```

-continued

---

 gtgaacgtgc acgctgacac tgggtgtgcc atagacatca gccggaaga gctgaga 2097

&lt;210&gt; SEQ ID NO 75

&lt;211&gt; LENGTH: 2097

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 75

tcatcaagcg gacacactcg actgttgaa tcatcacaaa cactcctaat tgttgtaat 60  
 tgtgtcacgc tcgacaaaga atcgtgcttc tagagctaca gattccttgg tcctcccggg 120  
 cggttctcgg cccgacagtt atacgatttt gcgccttacg gggcgacaaa caggaactaa 180  
 cctgaattcg ttttcttcgc tcccccgctc aacgattagg agtttcgttt acgttaatgg 240  
 tgttaggagt gacgtcagtg taaaacaaaa cgaagaccag tttttagtg acttcttaaa 300  
 atagttagtt gtacgtcacg tcaatcgttt ccgatagaat cagagactc ttgaccaacc 360  
 atatggtcac aatattgata tcttaattca ttatagttcc ttttattcac attacctgtt 420  
 ctacgattcc attttaacta ttttgttctt aatctattta tatttttacg acattgtctt 480  
 aacgtcaacg agtacgtttc gtgtggtggt tgtttgtag ctcggcttc tcttgatggt 540  
 tccaaatact taatatgtga gttgttacgg ttttttgggt tacattgtaa ttcgctcttt 600  
 tccttttctt ctaaagaacc aaaaaacaat ccacaacctc gacgttagcg gtcaccgcaa 660  
 cgacatagat tccaggacgt ggatcttccc cttcacttgt tctagtttcc acgagatgat 720  
 aggtgtttgt tccgacatca gtcgaatagt ttacctcaat cacagaattg gtcgtttcac 780  
 aatctggagt ttttgatata tctatttggt aacaatggat aacacttggt cgtttcgacg 840  
 tcgtatagtt tatactcttg acactatctc aagggtgttt tcttggtgtc tgatgatctc 900  
 taatgggtccc ttaaatacaca attacgtcca cattgatgtg gacattcgtg aatgtacaat 960  
 tgattatcac ttaataacag taattagtta ctatacggat attgtttact agtctttttc 1020  
 aattacaggt tgttacaagt ttatcaatct gtcgtttcaa tgagatagta caggattat 1080  
 tttctccttc agaatcgat acatcatggt aatggtgata taccacaata tctatgtggg 1140  
 acaacctttg atgtgtgtag gggagataca tgttggtgtg gttttcttcc cagggtgtag 1200  
 acaaattggt cttgactgtc tctaccatg acactgttac gtcctagtca tagaaagaag 1260  
 ggtgttcgac tttgtacatt tcaagttagt ttagctcata aaacactgtg ttacttgcca 1320  
 aattgtaatg gttcacttta tttagagacg ttacaactgt ataagttggg gtttatacta 1380  
 acattttaat actgaagttt ttgtctacat tcgtcgaggc aatagtgtag agatcctcgg 1440  
 taacacagta cgataccggt ttgatttaca tgcgtaggt tatttttagc accttagtat 1500  
 ttctgtaaaa gattgccccc gctaatacat agtttatttc cctacctgtg acacagacat 1560  
 ccattgtgta atataataca tttattcgtt cttccatttt cagagataca ttttccactt 1620  
 ggttattatt taaagatact gggtaatcat aaggggagac tacttaaaact acgtagtatt 1680  
 agagttcagt tgctcttcta attggtctcg gatcgtaaat aagcatttag gctacttaat 1740  
 aatgtattac atttacgacc atttaggtgg tgtttatagt actattgatg atattaatat 1800  
 cactaataatc attataacaa tagtaattaa cgacaacctg acgagaatat gacattccgg 1860  
 tcttcggtg gtcagtggtg ttcggttcta gttgactcac catatttatt ataacgtaaa 1920  
 tcattgttaa aactagacga gtttgaacgt ccgctacatc ttagtttagg acctgggagg 1980  
 gccctgtcca ggtatcgaga gtgcaaagag cgtcaacctc ctcaagacga gaaggagagg 2040

-continued

cacttgcaag tgcgactgtg acccacacgg tatctgtagt cggcggttct cgactct 2097

<210> SEQ ID NO 76  
 <211> LENGTH: 1327  
 <212> TYPE: PRT  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 76

Met Ser Lys Lys Pro Gly Gly Pro Gly Lys Ser Arg Ala Val Asn Met  
 1 5 10 15  
 Leu Lys Arg Gly Met Pro Arg Val Leu Ser Leu Ile Gly Leu Lys Gln  
 20 25 30  
 Lys Lys Arg Gly Gly Glu Leu Leu Ile Leu Lys Ala Asn Ala Ile Thr  
 35 40 45  
 Thr Ile Leu Thr Ala Val Thr Phe Cys Phe Ala Ser Gly Gln Asn Ile  
 50 55 60  
 Thr Glu Glu Phe Tyr Gln Ser Thr Cys Ser Ala Val Ser Lys Gly Tyr  
 65 70 75 80  
 Leu Ser Ala Leu Arg Thr Gly Trp Tyr Thr Ser Val Ile Thr Ile Glu  
 85 90 95  
 Leu Ser Asn Ile Lys Glu Asn Lys Cys Asn Gly Thr Asp Ala Lys Val  
 100 105 110  
 Lys Leu Ile Lys Gln Glu Leu Asp Lys Tyr Lys Asn Ala Val Thr Glu  
 115 120 125  
 Leu Gln Leu Leu Met Gln Ser Thr Pro Pro Thr Asn Asn Arg Ala Arg  
 130 135 140  
 Arg Glu Leu Pro Arg Phe Met Asn Tyr Thr Leu Asn Asn Ala Lys Lys  
 145 150 155 160  
 Thr Asn Val Thr Leu Ser Lys Lys Arg Lys Arg Arg Phe Leu Gly Phe  
 165 170 175  
 Leu Leu Gly Val Gly Ser Ala Ile Ala Ser Gly Val Ala Val Ser Lys  
 180 185 190  
 Val Leu His Leu Glu Gly Glu Val Asn Lys Ile Lys Ser Ala Leu Leu  
 195 200 205  
 Ser Thr Asn Lys Ala Val Val Ser Leu Ser Asn Gly Val Ser Val Leu  
 210 215 220  
 Thr Ser Lys Val Leu Asp Leu Lys Asn Tyr Ile Asp Lys Gln Leu Leu  
 225 230 235 240  
 Pro Ile Val Asn Lys Gln Ser Cys Ser Ile Ser Asn Ile Glu Thr Val  
 245 250 255  
 Ile Glu Phe Gln Gln Lys Asn Asn Arg Leu Leu Glu Ile Thr Arg Glu  
 260 265 270  
 Phe Ser Val Asn Ala Gly Val Thr Thr Pro Val Ser Thr Tyr Met Leu  
 275 280 285  
 Thr Asn Ser Glu Leu Leu Ser Leu Ile Asn Asp Met Pro Ile Thr Asn  
 290 295 300  
 Asp Gln Lys Lys Leu Met Ser Asn Asn Val Gln Ile Val Arg Gln Gln  
 305 310 315 320  
 Ser Tyr Ser Ile Met Ser Ile Ile Lys Glu Glu Val Leu Ala Tyr Val  
 325 330 335  
 Val Gln Leu Pro Leu Tyr Gly Val Ile Asp Thr Pro Cys Trp Lys Leu  
 340 345 350

-continued

---

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| His | Thr | Ser | Pro | Leu | Cys | Thr | Thr | Asn | Thr | Lys | Glu | Gly | Ser | Asn | Ile |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
| Cys | Leu | Thr | Arg | Thr | Asp | Arg | Gly | Trp | Tyr | Cys | Asp | Asn | Ala | Gly | Ser |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Val | Ser | Phe | Phe | Pro | Gln | Ala | Glu | Thr | Cys | Lys | Val | Gln | Ser | Asn | Arg |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |
| Val | Phe | Cys | Asp | Thr | Met | Asn | Ser | Leu | Thr | Leu | Pro | Ser | Glu | Ile | Asn |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Leu | Cys | Asn | Val | Asp | Ile | Phe | Asn | Pro | Lys | Tyr | Asp | Cys | Lys | Ile | Met |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |
| Thr | Ser | Lys | Thr | Asp | Val | Ser | Ser | Ser | Val | Ile | Thr | Ser | Leu | Gly | Ala |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |
| Ile | Val | Ser | Cys | Tyr | Gly | Lys | Thr | Lys | Cys | Thr | Ala | Ser | Asn | Lys | Asn |
| 450 |     |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Arg | Gly | Ile | Ile | Lys | Thr | Phe | Ser | Asn | Gly | Cys | Asp | Tyr | Val | Ser | Asn |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |
| Lys | Gly | Met | Asp | Thr | Val | Ser | Val | Gly | Asn | Thr | Leu | Tyr | Tyr | Val | Asn |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     |     | 495 |
| Lys | Gln | Glu | Gly | Lys | Ser | Leu | Tyr | Val | Lys | Gly | Glu | Pro | Ile | Ile | Asn |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |
| Phe | Tyr | Asp | Pro | Leu | Val | Phe | Pro | Ser | Asp | Glu | Phe | Asp | Ala | Ser | Ile |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Ser | Gln | Val | Asn | Glu | Lys | Ile | Asn | Gln | Ser | Leu | Ala | Phe | Ile | Arg | Lys |
| 530 |     |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Ser | Asp | Glu | Leu | Leu | His | Asn | Val | Asn | Ala | Gly | Lys | Ser | Thr | Thr | Asn |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Ile | Met | Ile | Thr | Thr | Ile | Ile | Ile | Val | Ile | Ile | Val | Ile | Leu | Leu | Ser |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     |     | 575 |
| Leu | Ile | Ala | Val | Gly | Leu | Leu | Leu | Tyr | Cys | Lys | Ala | Arg | Ser | Thr | Pro |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Val | Thr | Leu | Ser | Lys | Asp | Gln | Leu | Ser | Gly | Ile | Asn | Asn | Ile | Ala | Phe |
|     |     |     | 595 |     |     |     | 600 |     |     |     |     | 605 |     |     |     |
| Ser | Asn | Asn | Phe | Asp | Leu | Leu | Lys | Leu | Ala | Gly | Asp | Val | Glu | Ser | Asn |
| 610 |     |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |
| Pro | Gly | Pro | Gly | Gly | Lys | Thr | Gly | Ile | Ala | Val | Met | Ile | Gly | Leu | Ile |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |
| Ala | Cys | Val | Gly | Ala | Val | Thr | Leu | Ser | Asn | Phe | Gln | Gly | Lys | Val | Met |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     | 655 |     |
| Met | Thr | Val | Asn | Ala | Thr | Asp | Val | Thr | Asp | Val | Ile | Thr | Ile | Pro | Thr |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |
| Ala | Ala | Gly | Lys | Asn | Leu | Cys | Ile | Val | Arg | Ala | Met | Asp | Val | Gly | Tyr |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |
| Met | Cys | Asp | Asp | Thr | Ile | Thr | Tyr | Glu | Cys | Pro | Val | Leu | Ser | Ala | Gly |
| 690 |     |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |
| Asn | Asp | Pro | Glu | Asp | Ile | Asp | Cys | Trp | Cys | Thr | Lys | Ser | Ala | Val | Tyr |
| 705 |     |     |     |     | 710 |     |     |     |     |     | 715 |     |     |     | 720 |
| Val | Arg | Tyr | Gly | Arg | Cys | Thr | Lys | Thr | Arg | His | Ser | Arg | Arg | Ser | Arg |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |
| Arg | Ser | Leu | Thr | Val | Gln | Thr | His | Gly | Glu | Ser | Thr | Leu | Ala | Asn | Lys |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |
| Lys | Gly | Ala | Trp | Met | Asp | Ser | Thr | Lys | Ala | Thr | Arg | Tyr | Leu | Val | Lys |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |
| Thr | Glu | Ser | Trp | Ile | Leu | Arg | Asn | Pro | Gly | Tyr | Ala | Leu | Val | Ala | Ala |

-continued

| 770  |     |     |     |     | 775 |      |      |     |     | 780 |     |      |      |     |     |
|------|-----|-----|-----|-----|-----|------|------|-----|-----|-----|-----|------|------|-----|-----|
| Val  | Ile | Gly | Trp | Met | Leu | Gly  | Ser  | Asn | Thr | Met | Gln | Arg  | Val  | Val | Phe |
| 785  |     |     |     |     | 790 |      |      |     |     | 795 |     |      |      |     | 800 |
| Val  | Val | Leu | Leu | Leu | Leu | Val  | Ala  | Pro | Ala | Tyr | Ser | Phe  | Asn  | Cys | Leu |
|      |     |     |     | 805 |     |      |      |     | 810 |     |     |      |      | 815 |     |
| Gly  | Met | Ser | Asn | Arg | Asp | Phe  | Leu  | Glu | Gly | Val | Ser | Gly  | Ala  | Thr | Trp |
|      |     |     | 820 |     |     |      |      | 825 |     |     |     |      | 830  |     |     |
| Val  | Asp | Leu | Val | Leu | Glu | Gly  | Asp  | Ser | Cys | Val | Thr | Ile  | Met  | Ser | Lys |
|      | 835 |     |     |     |     |      | 840  |     |     |     |     | 845  |      |     |     |
| Asp  | Lys | Pro | Thr | Ile | Asp | Val  | Lys  | Met | Met | Asn | Met | Glu  | Ala  | Ala | Asn |
| 850  |     |     |     |     |     | 855  |      |     |     |     | 860 |      |      |     |     |
| Leu  | Ala | Glu | Val | Arg | Ser | Tyr  | Cys  | Tyr | Leu | Ala | Thr | Val  | Ser  | Asp | Leu |
| 865  |     |     |     | 870 |     |      |      |     |     | 875 |     |      |      |     | 880 |
| Ser  | Thr | Lys | Ala | Ala | Cys | Pro  | Ala  | Met | Gly | Glu | Ala | His  | Asn  | Asp | Lys |
|      |     |     |     | 885 |     |      |      |     | 890 |     |     |      |      |     | 895 |
| Arg  | Ala | Asp | Pro | Ala | Phe | Val  | Cys  | Arg | Gln | Gly | Val | Val  | Asp  | Arg | Gly |
|      |     |     | 900 |     |     |      |      | 905 |     |     |     |      |      | 910 |     |
| Trp  | Gly | Asn | Gly | Cys | Gly | Leu  | Phe  | Gly | Lys | Gly | Ser | Ile  | Asp  | Thr | Cys |
|      |     | 915 |     |     |     |      | 920  |     |     |     |     |      | 925  |     |     |
| Ala  | Lys | Phe | Ala | Cys | Ser | Thr  | Lys  | Ala | Ile | Gly | Arg | Thr  | Ile  | Leu | Lys |
| 930  |     |     |     |     |     | 935  |      |     |     |     | 940 |      |      |     |     |
| Glu  | Asn | Ile | Lys | Tyr | Glu | Val  | Ala  | Ile | Phe | Val | His | Gly  | Pro  | Thr | Thr |
| 945  |     |     |     |     | 950 |      |      |     |     |     | 955 |      |      |     | 960 |
| Val  | Glu | Ser | His | Gly | Asn | Tyr  | Ser  | Thr | Gln | Val | Gly | Ala  | Thr  | Gln | Ala |
|      |     |     |     | 965 |     |      |      |     | 970 |     |     |      |      |     | 975 |
| Gly  | Arg | Phe | Ser | Ile | Thr | Pro  | Ala  | Ala | Pro | Ser | Tyr | Thr  | Leu  | Lys | Leu |
|      |     |     | 980 |     |     |      |      | 985 |     |     |     |      |      |     | 990 |
| Gly  | Glu | Tyr | Gly | Glu | Val | Thr  | Val  | Asp | Cys | Glu | Pro | Arg  | Ser  | Gly | Ile |
|      |     | 995 |     |     |     |      | 1000 |     |     |     |     |      | 1005 |     |     |
| Asp  | Thr | Asn | Ala | Tyr | Tyr | Val  | Met  | Thr | Val | Gly | Thr | Lys  | Thr  | Phe |     |
| 1010 |     |     |     |     |     | 1015 |      |     |     |     |     | 1020 |      |     |     |
| Leu  | Val | His | Arg | Glu | Trp | Phe  | Met  | Asp | Leu | Asn | Leu | Pro  | Trp  | Ser |     |
| 1025 |     |     |     |     |     | 1030 |      |     |     |     |     | 1035 |      |     |     |
| Ser  | Ala | Gly | Ser | Thr | Val | Trp  | Arg  | Asn | Arg | Glu | Thr | Leu  | Met  | Glu |     |
| 1040 |     |     |     |     |     | 1045 |      |     |     |     |     | 1050 |      |     |     |
| Phe  | Glu | Glu | Pro | His | Ala | Thr  | Lys  | Gln | Ser | Val | Ile | Ala  | Leu  | Gly |     |
| 1055 |     |     |     |     |     | 1060 |      |     |     |     |     | 1065 |      |     |     |
| Ser  | Gln | Glu | Gly | Ala | Leu | His  | Gln  | Ala | Leu | Ala | Gly | Ala  | Ile  | Pro |     |
| 1070 |     |     |     |     |     | 1075 |      |     |     |     |     | 1080 |      |     |     |
| Val  | Glu | Phe | Ser | Ser | Asn | Thr  | Val  | Lys | Leu | Thr | Ser | Gly  | His  | Leu |     |
| 1085 |     |     |     |     |     | 1090 |      |     |     |     |     | 1095 |      |     |     |
| Lys  | Cys | Arg | Val | Lys | Met | Glu  | Lys  | Leu | Gln | Leu | Lys | Gly  | Thr  | Thr |     |
| 1100 |     |     |     |     |     | 1105 |      |     |     |     |     | 1110 |      |     |     |
| Tyr  | Gly | Val | Cys | Ser | Lys | Ala  | Phe  | Lys | Phe | Leu | Gly | Thr  | Pro  | Ala |     |
| 1115 |     |     |     |     |     | 1120 |      |     |     |     |     | 1125 |      |     |     |
| Asp  | Thr | Gly | His | Gly | Thr | Val  | Val  | Leu | Glu | Leu | Gln | Tyr  | Thr  | Gly |     |
| 1130 |     |     |     |     |     | 1135 |      |     |     |     |     | 1140 |      |     |     |
| Thr  | Asp | Gly | Pro | Cys | Lys | Val  | Pro  | Ile | Ser | Ser | Val | Ala  | Ser  | Leu |     |
| 1145 |     |     |     |     |     | 1150 |      |     |     |     |     | 1155 |      |     |     |
| Asn  | Asp | Leu | Thr | Pro | Val | Gly  | Arg  | Leu | Val | Thr | Val | Asn  | Pro  | Phe |     |
| 1160 |     |     |     |     |     | 1165 |      |     |     |     |     | 1170 |      |     |     |
| Val  | Ser | Val | Ala | Thr | Ala | Asn  | Ala  | Lys | Val | Leu | Ile | Glu  | Leu  | Glu |     |
| 1175 |     |     |     |     |     | 1180 |      |     |     |     |     | 1185 |      |     |     |

-continued

Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Glu Gln  
 1190 1195 1200  
 Gln Ile Asn His His Trp His Lys Ser Gly Ser Ser Ile Gly Lys  
 1205 1210 1215  
 Ala Phe Thr Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu  
 1220 1225 1230  
 Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly Gly Val Phe Thr  
 1235 1240 1245  
 Ser Val Gly Lys Ala Val His Gln Val Phe Gly Gly Ala Phe Arg  
 1250 1255 1260  
 Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu Gly  
 1265 1270 1275  
 Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg Ser Ile  
 1280 1285 1290  
 Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser  
 1295 1300 1305  
 Val Asn Val His Ala Asp Thr Gly Cys Ala Ile Asp Ile Ser Arg  
 1310 1315 1320  
 Gln Glu Leu Arg  
 1325

&lt;210&gt; SEQ ID NO 77

&lt;211&gt; LENGTH: 4100

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Construct

&lt;400&gt; SEQUENCE: 77

```

gatcctaata cgactcacta tagagtagtt cgcctgtgtg agctgacaaa cttagtagtg 60
tttgtgagga ttaacaacaa ttaacacagt gcgagctggt tcttagcacg aagatctcga 120
tgtctaagaa accaggaggg cccggcaaga gccgggctgt caatatgcta aaacgcggaa 180
tgccccgcgt gttgtccttg attggactta agcaaaaagaa gcgagggggc gagttgctaa 240
tcctcaaagc aaatgcaatt accacaatcc tcaactgcagt cacattttgt tttgcttctg 300
gtcaaaacat cactgaagaa ttttatcaat caacatgcag tgcagttagc aaaggctatc 360
ttagtgtctc gagaactggt tggttatacca gtgttataac tatagaatta agtaatatca 420
aggaaaataa gtgtaatgga acagatgcta aggtaaaatt gataaaacaa gaattagata 480
aatataaaaa tgctgtaaca gaattgcagt tgctcatgca aagcacacca ccaacaaaca 540
atcgagccag aagagaacta ccaaggttta tgaattatac actcaacaat gccaaaaaaa 600
ccaatgtaac attaagcaag aaaaggaaaa gaagatttct tggttttttg ttaggtggtg 660
gatctgcaat cgccagtggc gttgctgtat ctaaggtoct gcacctagaa ggggaagtga 720
acaagatcaa aagtgtctca ctatccacaa acaaggctgt agtcagetta tcaaatggag 780
ttagtgtctt aaccagcaaa gtgtagacc tcaaaaacta tatagataaa caattgttac 840
ctattgtgaa caagcaaacg tgcagcatat caaatataga aactgtgata gagttccaac 900
aaaagaacaa cagactacta gagattacca gggaatttag tgттаatgca ggtgtaacta 960
cacctgtaag cacttacatg ttaactaata gtgaattatt gtcattaatc aatgatatgc 1020
ctataacaaa tgatcagaaa aagttaatgt ccaacaatgt tcaaatagtt agacagcaaa 1080
gttactctat catgtccata ataaaagagg aagtcttagc atatgtagta caattaccac 1140

```

-continued

---

|             |             |             |            |            |            |      |
|-------------|-------------|-------------|------------|------------|------------|------|
| tatatggtgt  | tatagataca  | cctgttga    | aactacacac | atcccccta  | tgtacaacca | 1200 |
| acacaaaaga  | agggtccaac  | atctgttta   | caagaactga | cagaggatgg | tactgtgaca | 1260 |
| atgcaggatc  | agtatctttc  | ttcccacaag  | ctgaaacatg | taaagttcaa | tcaaatcgag | 1320 |
| tattttgtga  | cacaatgaac  | agtttaacat  | taccaagtga | aataaatctc | tgcaatggtg | 1380 |
| acatattcaa  | ccccaaatat  | gattgtaaaa  | ttatgacttc | aaaaacagat | gtaagcagct | 1440 |
| cogttatcac  | atctctagga  | gccattgtgt  | catgctatgg | caaaactaaa | tgtacagcat | 1500 |
| ccaataaaaa  | tcgtggaatc  | ataaagacat  | tttctaacgg | gtgcgattat | gtatcaata  | 1560 |
| aagggatgga  | cactgtgtct  | gtaggtaaca  | cattatatta | tgtaataaag | caagaaggta | 1620 |
| aaagtctcta  | tgtaaaaggt  | gaaccaataa  | taaatttcta | tgaccatta  | gtattccctc | 1680 |
| ctgatgaatt  | tgatgcatca  | atatctcaag  | tcaacgagaa | gattaaccag | agcctagcat | 1740 |
| ttattcgtaa  | atccgatgaa  | ttattacata  | atgtaaatgc | tggtaaatcc | accacaaata | 1800 |
| tcatgataac  | tactataaatt | atagtgatta  | tagtaatatt | gttatcatta | attgctgttg | 1860 |
| gactgctctt  | atactgtaag  | gccagaagca  | caccagtcac | actaagcaaa | gatcaactga | 1920 |
| gtggtataaa  | taatattgca  | tttagtaaca  | attttgatct | gctcaaaact | gcaggcgatg | 1980 |
| tagaatcaaa  | tcctggaccc  | ggaggaaaga  | coggtattgc | agtcatgatt | ggcctgatcg | 2040 |
| cctgcgtagg  | agcagttacc  | ctctctaact  | tccaagggaa | ggtgatgatg | acggtaaatg | 2100 |
| ctactgacgt  | cacagatgtc  | atcacgattc  | caacagctgc | tggaaagaac | ctatgcattg | 2160 |
| tcagagcaat  | ggatgtggga  | tacatgtgcg  | atgatactat | cacttatgaa | tgcccagtgc | 2220 |
| tgtcggctgg  | taatgatcca  | gaagacatcg  | actgttggtg | cacaaagtca | gcagtctacg | 2280 |
| tcaggatggt  | aagatgcacc  | aagacacgcc  | actcaagacg | cagtcggagg | tcactgacag | 2340 |
| tgacagacaca | cggagaaagc  | actctagcga  | acaagaaggg | ggcttggatg | gacagcacca | 2400 |
| aggccacaag  | gtatttggtg  | aaaacagaat  | catggatctt | gaggaacctc | ggatatgcc  | 2460 |
| tgggtggcagc | cgtcatgggt  | tggatgcttg  | ggagcaacac | catgcagaga | gttgtgtttg | 2520 |
| tcgtgctatt  | gcttttgggt  | gccccagctt  | acagctttaa | ctgccttggg | atgagcaaca | 2580 |
| gagacttctt  | ggaaggagtg  | tctggagcaa  | catgggtgga | tttggttctc | gaaggcgaca | 2640 |
| gctgcgtgac  | tatcatgtct  | aaggacaagc  | ctaccatcga | tgtgaagatg | atgaatatgg | 2700 |
| aggcgcccaa  | cctggcagag  | gtccgcagtt  | attgctattt | ggctaccgtc | agcgtctctc | 2760 |
| ccaccaaagc  | tgcgtgcccg  | gccatgggag  | aagctcacia | tgacaaacgt | gctgacccag | 2820 |
| cttttgtgtg  | cagacaagga  | gtgggtggaca | ggggctgggg | caacggctgc | ggactatttg | 2880 |
| gcaaaggaag  | cattgacaca  | tgcgccaaat  | ttgcctgctc | taccaaggca | ataggaagaa | 2940 |
| ccattttgaa  | agagaatatc  | aagtacgaag  | tggccatttt | tgtccatgga | ccaactactg | 3000 |
| tggagtgcga  | cggaaactac  | tccacacagg  | ttggagccac | tcaggcaggg | agattcagca | 3060 |
| tcactcctgc  | ggcgcttca   | tacacactaa  | agcttgagga | atatggagag | gtgacagtgg | 3120 |
| actgtgaacc  | acggtcaggg  | attgacacca  | atgcatacta | cgtgatgact | gttggaacaa | 3180 |
| agacgttctt  | ggtccatcgt  | gagtggttca  | tggacctcaa | cctcccttgg | agcagtgctg | 3240 |
| gaagtactgt  | gtggaggaac  | agagagacgt  | taatggagt  | tgaggaacca | cacgccacga | 3300 |
| agcagtctgt  | gatagcattg  | ggctcacaag  | agggagctct | gcatcaagct | ttggctggag | 3360 |
| ccattcctgt  | ggaattttca  | agcaaacctg  | tcaagttgac | gtcgggtcat | ttgaagtgtg | 3420 |
| gagtgaagat  | ggaaaaattg  | cagttgaagg  | gaacaaccta | tggcgtctgt | tcaaaggctt | 3480 |
| tcaagttctc  | tgggactccc  | gcagacacag  | gtcacggcac | tgtggtgttg | gaattgcagt | 3540 |

-continued

---

```

aactggcac ggtggaact tgcaaagttc ctatctctgc agtggcttca ttgaacgacc 3600
taacgccagt gggcagattg gtcactgtca acccttttgt ttcagtggcc acggccaacg 3660
ctaaggctct gattgaattg gaaccaccct ttggagactc atacatagtg gtgggcagag 3720
gagaacaaca gatcaatcac cactggcaca agtctggaag cagcattggc aaagccttta 3780
caaccaccct caaaggagcg cagagactag ccgctctagg agacacagct tgggactttg 3840
gatcagttgg aggggtgttc acctcagttg ggaaggctgt ccatcaagtg ttcggaggag 3900
cattccgctc actgttcgga ggcagtctct ggataacgca aggattgctg ggggctctcc 3960
tgttggtgat gggcatcaat gctcgtgaca ggtccatagc tctcacgttt ctgcagttg 4020
gaggagttct gctcttctc tccgtgaacg tgcacgctga cactgggtgt gccatagaca 4080
tcagccggca agagctgaga 4100

```

```

<210> SEQ ID NO 78
<211> LENGTH: 4100
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 78

```

```

ctaggattat gctgagtgat atctcatcaa gggacacac tcgactgttt gaatcatcac 60
aaacactcct aattgttgtt aattgtgtca cgctcgacaa agaatcgtgc ttctagagct 120
acagattctt tggctctccc gggccgttct cggcccgaca gttatacgat tttgcgctt 180
acggggcgca caacaggaac taacctgaat tcgttttctt cgctcccccg ctcaacgatt 240
aggagtttcc tttacgttaa tgggttagg agtgacgtca gtgtaaaaca aaacgaagac 300
cagttttgta gtgactctct aaaatagtta gttgtacgtc acgtcaatcg tttccgatag 360
aatcacgaga ctcttgacca accataggt cacaaatattg atatcttaac tcattatagt 420
tcccttttatt cacattaact tgtctacgat tccattttaa ctattttgtt cttaatctat 480
ttatattttt acgacattgt cttaacgtca acgagtaact ttcgtgtggt ggttgtttgt 540
tagctcggtc ttctcttgat ggttccaaat acttaaatatg tgagttgtta cggttttttt 600
ggttacattg taattcgttc ttttctttt ctcttaaaga accaaaaaac aatccacaac 660
ctagacgtta gcggtcaccg caacgacata gattccagga cgtggatctt ccccttccact 720
tgttctagtt ttcacgagat gataggtgtt tgttccgaca tcagtcgaat agtttacctc 780
aatcacagaa ttggtcgttt cacaactctgg agtttttgat atatctattt gttaacaatg 840
gataacactt gttcgtttcc acgtcgtata gtttatatct ttgacactat ctcaaggttg 900
ttttcttggt gtctgatgat ctctaattgt ccccttaaatc acaattacgt ccacattgat 960
gtggacattc gtgaatgtac aattgattat cacttaataa cagtaattag ttactatacg 1020
gatattgttt actagctttt ttcaattaca gtttgttaca agtttatcaa tctgtcgttt 1080
caatgagata gtacaggat tttttctccc ttcagaatcg tatacatcat gttaatggtg 1140
atataccaca atatctatgt gggacaacct ttgatgtgtg taggggagat acatgttggg 1200
tgtgttttct tcccaggttg tagacaaatt gttcttgact gtctcctacc atgacactgt 1260
tacgccttag tcatagaaag aagggtgttc gactttgtac atttcaagtt agtttagctc 1320
ataaaacact gtgttacttg tcaaatgtga atggttccact ttatttagag acgttacaac 1380
tgtataagtt ggggtttata ctaacatttt aatactgaag tttttgtcta cattcgtcga 1440

```

-continued

---

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ggcaatagtg tagagatcct cgtaacaca gtacgatacc gttttgattt acatgtcgta    | 1500 |
| ggttatTTTT agcaccttag tatttctgta aaagattgcc cacgctaata catagtttat   | 1560 |
| ttccctacct gtgacacaga catccattgt gtaatataat acattttattc gttcttccat  | 1620 |
| tttcagagat acattttoca cttggttatt atttaaagat actgggtaat cataagggga   | 1680 |
| gactacttaa actacgtagt tatagagttc agttgctctt ctaattggtc tcggatcgta   | 1740 |
| aataagcatt taggctaact aataatgtat tacatttacg accatttagg tgggttttat   | 1800 |
| agtactattg atgatattaa tatcactaat atcattataa caatagtaat taacgacaac   | 1860 |
| ctgacgagaa tatgacattc cggctctcgt gtggtcagtg tgattcgttt ctagttgact   | 1920 |
| caccatattt attataacgt aaatcattgt taaaactaga cgagtttgaa cgtccgctac   | 1980 |
| atcttagttt aggacctggg cctcctttct ggccataacg tcagtactaa cgggactagc   | 2040 |
| ggacgcatcc tcgtcaatgg gagagattga aggttcocct ccaactactac tgccatttac  | 2100 |
| gatgactgca gtgtctacag tagtgctaag gttgtcgacg acctttcttg gatacgtaac   | 2160 |
| agtctcgta cctacacct atgtacacgc tactatgata gtgaatactt acgggtcacg     | 2220 |
| acagccgacc attactaggt cttctgtagc tgacaaccac gtgtttcagt cgtcagatgc   | 2280 |
| agtccatacc ttctacgtgg ttctgtcggg tgagttctgc gtcagcctcc agtgactgtc   | 2340 |
| acgtctgtgt gcctctttcg tgagatcgtt tgttcttccc ccgaacctac ctgtcgtggg   | 2400 |
| tcgggtgttc cataaaccat ttttgtctta gtacctagaa ctccctggga cctatacggg   | 2460 |
| accaccgtcg gcagtaacca acctacgaac cctcgttggt gtacgtctct caacacaaa    | 2520 |
| agcacgataa cgaaaaccac cggggctgaa tgtcgaaatt gacggaacct tactcgttgt   | 2580 |
| ctctgaagaa ccttcctcac agacctcgtt gtacctacct aaaccaagag cttccgctgt   | 2640 |
| cgacgcactg atagtacaga ttctgttctg gatggtagct acacttctac tacttatacc   | 2700 |
| tccgccggtt ggacctctc caggcgtcaa taacgataaa ccgatggcag tcgctagaga    | 2760 |
| ggtggtttcg acgcaacggc cgttacacct ttcgagtggt actgtttgea cgactgggtc   | 2820 |
| gaaaacacac gtctgttctc caccacctgt ccccgacccc gttgccgacg cctgataaac   | 2880 |
| cgtttccttc gtaactgtgt acgcggttta aacggacgag atggttccgt tctcctctt    | 2940 |
| ggtaaaaact tctcttatag ttcattcttc accggtaaaa acaggtacct ggttgatgac   | 3000 |
| acctcagcgt gcctttgatg aggtgtgtcc aacctcggtg agtccgtccc tctaagtcgt   | 3060 |
| agtgaggacg ccgcggaagt atgtgtgatt tcgaaacctc tataacctctc cactgtcacc  | 3120 |
| tgacacttgg tgccagtcct taactgtggt tacgtatgat gcactactga caaccttgtt   | 3180 |
| tctgcaagaa ccaggtagca ctccaccaagt acctggagtt ggaggggaacc tcgtcacgac | 3240 |
| cttcatgaca cacctccttg tctctctgca attacctcaa actccttggg gtgoggtgct   | 3300 |
| tcgtcagaca ctatcgtaac ccgagtggtc tccctcgaga cgtagttcga aaccgacctc   | 3360 |
| ggtaaggaca ccttaaaagt tcgttgtgac agttcaactg cagcccagta aacttcacat   | 3420 |
| ctcacttcta cctttttaac gtcaacttcc cttgttgat accgcagaca agtttccgaa    | 3480 |
| agttcaaaga accctgaggg cgtctgtgtc cagtgccgtg acaccacaac cttaacgtca   | 3540 |
| tgtgaccgtg cctacctgga acgtttcaag gatagagcag tcaccgaagt aacttctgtg   | 3600 |
| attgctgtca cccgtctaac cagtgtgacgt tgggaaaaa aagtcaccgg tgccggttgc   | 3660 |
| gattccagga ctaacttaac cttgggtggga aacctctgag tatgtatcac caccctctc   | 3720 |
| ctcttgtgt ctagttagtg gtgacctgtt tcagacctc gtcgtaaccg tttcggaaat     | 3780 |
| gttggtggga gtttcctcgc gtctctgac gccgagatcc tctgtgtcga acctgaaac     | 3840 |

-continued

---

```

ctagtcaccc tccccacaag tggagtcaac ccttcgcaca ggtagttcac aagcctcctc 3900
gtaaggcgag tgacaagcct ccgtacagga cctattgcgt tctaacgcac ccccgagagg 3960
acaacaccta cccgtagtta cgagcactgt ccaggtatcg agagtgcata gagcgtcaac 4020
ctcctcaaga cgagaaggag aggcacttgc acgtgcgact gtgaccacac cggtatctgt 4080
agtcggccgt tctcgactct 4100

```

---

What is claimed is:

1. A replication-deficient pseudoinfectious flavivirus comprising a West Nile virus genome comprising (i) one or more deletions in nucleotide sequences encoding capsid (C) protein, and deletion of the nucleotide sequences encoding pre-membrane (prM) and/or envelope (E) proteins, and (ii) a sequence encoding a heterologous immunogen inserted in place of or in combination with said deletion of the prM and/or E nucleotide sequences,

wherein the heterologous immunogen is from a pathogen selected from the group consisting of a rabies virus, influenza virus, respiratory syncytial virus (RSV), and human immunodeficiency virus (HIV).

2. The replication-deficient pseudoinfectious flavivirus of claim 1, wherein said heterologous immunogen is a rabies virus G protein.

3. The replication-deficient pseudoinfectious flavivirus of claim 1, wherein said heterologous immunogen is an influenza virus antigen selected from the group consisting of M2, hemagglutinin (HA), and neuraminidase (NA), or an immunogenic fragment thereof.

4. The replication-deficient pseudoinfectious flavivirus of claim 1, wherein said heterologous immunogen is a codon-optimized HIV protein or an immunogenic fragment thereof selected from the group consisting of gag, env, tat/nef, gp120, and gp160.

5. The replication-deficient pseudoinfectious flavivirus of claim 1, wherein said genome is packaged in a particle comprising pre-membrane (prM) and envelope (E) sequences from a flavivirus that is the same or different from that of the genome.

6. A method of inducing an immune response to an immunogen in a subject, the method comprising administering to the subject one or more replication-deficient pseudoinfectious flaviviruses of claim 1.

7. The replication-defective pseudoinfectious flavivirus of claim 1, comprising multiple heterologous immunogens.

8. The replication-deficient pseudoinfectious flavivirus of claim 1, wherein said heterologous immunogen is RSV F protein or an immunogenic fragment thereof.

\* \* \* \* \*

UNITED STATES PATENT AND TRADEMARK OFFICE  
**CERTIFICATE OF CORRECTION**

PATENT NO. : 9,217,158 B2  
APPLICATION NO. : 13/633436  
DATED : December 22, 2015  
INVENTOR(S) : Pugachev et al.

Page 1 of 1

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

On the Title Page:

The first or sole Notice should read --

Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b)  
by 181 days.

Signed and Sealed this  
Twelfth Day of July, 2016



Michelle K. Lee  
*Director of the United States Patent and Trademark Office*